Molecular detection methods for the investigation of potential sources of Campylobacter infection

Thesis

UNRESTRICTED

Molecular Detection Methods for the Investigation of Potential Sources of Campylobacter Infection

Emma Louisa Best BSc (hons) MSc

A thesis submitted in partial fulfilment of the requirements of the Open University for the degree of Doctor of Philosophy

August 2004

Campylobacter Reference Unit, Laboratory of Enteric Pathogens, Specialist and Reference Microbiology Division, Health Protection Agency Colindale, London. NW9 5HT.

Health Protection Agency North West Laboratory, Manchester Medical Microbiology Partnership, Manchester Royal Infirmary, Manchester. M13 9WL.

Submission date: 10 August 2004
Award date: 18 October 2004

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Declaration

The contents of this thesis are the result of the independent work of Emma Best, and have not been submitted in whole or part for any other academic award at the Open University or any other institution. All contributions from others are indicated.

Signed **Date** 12/10/04.

Emma Best

Abstract

The rapid detection and identification of *Campylobacter jejuni* isolates to probable strain level would significantly inform the epidemiological investigation of *C.jejuni* infection. At the outset of this project the molecular fingerprinting techniques PFGE and AFLP were proven to be equally discriminatory for identification of outbreak strains of campylobacter, however both techniques were time consuming and not directly applicable to potential sources of infection. Real time PCR approaches were pursued for the purpose of developing methods, which would be specific and robust for the detection of specific strains of campylobacter. A duplex real-time polymerase chain reaction (PCR) assay for speciation of *Campylobacter jejuni* and *C.coli* using real time platforms was developed. This enabled a turnaround time of three hours and was applied for direct speciation from sources of infection including meat samples. The development of real time PCR assays for allelic discrimination of strain associated single nucleotide polymorphisms (SNPs) based upon MLST locus alleles offered a possible approach for rapid strain detection. Single nucleotide polymorphisms defining key alleles diagnostic for six major clonal complexes were identified, following a detailed analysis of the available MLST data. Allelic discrimination assays based on real time PCR systems were designed to detect the SNPs and be specific for clonal complexes ST-21, ST-45, ST-48, ST-61 ST-206 and ST-257. SNP based assays were evaluated using panels of isolates from human infections, poultry, the environment, and the MLST reference collection, which had previously been characterized by MLST. Real time allelic discrimination assays allowed the rapid detection of *C.jejuni* isolates and preliminary strain identification directly from foods and environmental specimens. The ability to combine detection with the identification of epidemiologically important information beyond genus or species identification represents a major new concept in the use of nucleic acid amplification techniques for the improved detection of pathogens particularly pathogens of major public health importance such as *C.jejuni*.

List of Contents **Page**

Chapter 2 Materials and Methods

List of Tables **Page**

List of Figures **Page**

Abbreviations

AFLP	Amplified Fragment Length Polymorphism
BLAST	Basic local alignment search tool
bp	Base pairs (DNA)
CBA	Columbia Blood Agar
CCDA	Cefalozin charcoal deoxycholate agar
CFU	Colony forming unit
CPHL	Central Public Health Laboratory
CRU	Campylobacter Reference Unit
C_T	Threshold Cycle
dsDNA	double stranded DNA
DNA	Deoxyribonucleic Acid
dNTP	deoxynucleotide triphosphate
EDTA	Ethylene diamine tetra acetic acid
ELISA	Enzyme linked immunosorbent assay
HS	Heat stable serotype
HPA	Health Protection Agency
kb	kilobase
LEP	Laboratory of Enteric Pathogens
MF	McFarland Standard
MLST	Multi Locus Sequence Typing
MLEE	Multi Locus Enzyme Electrophoresis
MGB	Minor Groove Binder
ml	Millilitre
NTC	Non template control
NCTC	National Collection of Type Cultures, London. UK
PCR	Polymerase Chain reaction
PT	Phage Type
PFGE	Pulsed Field Gel Electrophoresis
RDNC	Reacts with phage but does not conform to specific type
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
rpm	revolutions per minute
SNP	Single Nucleotide Polymorphism
ST	Sequence Type
Taq	*Taq* DNA polymerase
T_m	Melting temperature
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
UT	untypeable
µg	microgram
µl	microlitre
µM	micromolar

Acknowledgements

I wish to thank my supervisors, Professor Andrew Fox, Professor Eric Bolton and Mrs Jenny Frost for initially choosing me to do this project and for all their help, guidance, sharing of knowledge, support, enthusiasm and experience over the last three years. Also I wish to thank the Health Protection Agency for funding through a PhD studentship.

Thank you to all my friends and colleagues from the Campylobacter and Helicobacter Reference Unit (formerly the Campylobacter Reference Unit), HPA Colindale and the Molecular Epidemiology Unit, Manchester HPA who have helped directly or indirectly with this project. I also wish to acknowledge Dr Kathie Grant in this study for her help, advice and useful discussions, Dr. Henry Smith for proof reading papers for publication and Ankur Agrawal for his bio informatics expertise.

I wish to thank Frances Colles, Roisin Ure and Martin Maiden, Department of Zoology at the University of Oxford for all their help in setting up of MLST and use of sequencing facilities. This thesis made use of the Campylobacter Multi Locus Sequence Typing website (http://pubmlst.org/campylobacter/) developed by Dr Man-Suen Chan and Dr. Keith Jolley and sited at the University of Oxford. Initial development of this site was funded by the Wellcome Trust, maintenance is funded by DEFRA.

Finally I wish to express my utmost thanks to my family for their unfailing support and encouragement throughout this study and to Alex for his help with proof reading and for keeping me relatively sane throughout the more difficult times.

Publications and presentations arising from this work

Champion OL, **Best EL**, Frost JA (2002) Comparison of Pulsed Field Gel Electrophoresis and Amplified Length Polymorphism techniques for investigating outbreaks of enteritis due to campylobacter. *Journal of Clinical Microbiology* **40:** 2263 – 2265

Best EL, Powell EJ, Swift C, Grant KA, Frost JA. (2003) Applicability of a rapid duplex real-time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from charcoal swabs and culture plates. *FEMS Microbiology Letters* **229** (2) 237-241

Best EL, Fox AJ, Frost JA , Bolton FJ. (2004) Identification of *Campylobacter jejuni* Multi Locus Sequence Typing ST-21 Clonal Complex by SNP Analysis. *Journal of Clinical Microbiology* **42** (6) 2836-2839.

Best EL, Fox AJ, Frost JA , Bolton FJ. Real Time strain characterisation of *Campylobacter jejuni* by detection of single nucleotide polymorphisms (SNPs) predictive for six MLST clonal complexes using Taqman Technology. *Journal of Clinical Microbiology* **submitted**

Best EL, Champion O, Richardson JF, Frost JA. (2001) A comparison of the molecular fingerprinting techniques Pulsed Field Gel Electrophoresis and Amplified Fragment Length Polymorphism in campylobacter outbreak investigations. International Workshop On Campylobacter, Helicobacter And Related Organisms, Freiburg, Germany *International Journal of Medical Microbiology* **291;** Suppl 31:

Best EL, Powell E, Swift C, Grant K, Frost JA (2003) Real-Time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from culture plates, charcoal swabs and spiked food samples. International Workshop on Campylobacter Helicobacter and Related Organisms, Aarhus, Denmark *International Journal of Medical Microbiology* **293;** Suppl 55

Best EL, Fox A, Frost JA, Bolton FJ (2003) *Campylobacter jejuni* MLST clonal complex identification by SNP analysis using melting peak analysis in the Lightcycler. International Workshop on Campylobacter Helicobacter and Related Organisms, Aarhus, Denmark. *International Journal of Medical Microbiology* **293;** Suppl 55

Best EL, Powell E, Swift C, Frost JA (2003) Applicability of a rapid duplex real time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from culture, charcoal swabs and spiked food samples. *HPA Annual Scientific Meeting.* Warwick University. UK.

Best EL, Fox A.J. Frost JA, Bolton FJ (2004) Preliminary strain characterisation of *Campylobacter jejuni*, by real time detection of single nucleotide polymorphisms (SNPs) within clonal complexes. *International Conference on Emerging Infectious Diseases.* Atlanta. USA.

Chapter 1

Introduction

Molecular Detection Methods for the Investigation of Potential Sources of Campylobacter Infection

Chapter 1. Introduction

1.1 Introduction to the investigation

The reduction in foodborne gastrointestinal infection by 20% by 2006 is a major target for the Food Standards Agency (www.foodstandards.gov.uk). Key to achieving this target will be a reduction in the burden of infection due to *Campylobacter*. Lack of routinely used methods for detection and subtyping (Frost, 2001) of campylobacter has sustained the epidemiological paradox of campylobacter infection (Cowden, 1992), thus supporting the realistic hypothesis that a substantial proportion of sporadic campylobacter infections go undetected. Additionally the link between a campylobacter infection and exposure to a specific source is rarely established, due to an absence of timely molecular methods for strain characterisation.

The proportion of suspect potential sources likely to be Campylobacter-positive is high in comparison to those that may be salmonella-positive. An individual food sample can also carry more than one campylobacter strain, where the strain responsible for human infection may be the minority component. Campylobacter has a relatively long incubation period compared to other enteric pathogens and this, together with the complexity of isolation and current methods for phenotyping (Frost *et al.*, 1998; Frost *et al.*, 1999) and genotyping (Desai *et al.*, 2001), can make the detection of specific strains within potential sources a time consuming and difficult

process. Although current methods have defined epidemiological groupings they are unable to provide the timely fingerprinting of isolates for robust early detection and case cluster recognition.

Campylobacters vary in their potential to cause human disease however they do not possess the recognised virulence mechanisms as found in other organisms. This was also confirmed through genome analysis of the sequenced strain *C. jejuni* NCTC11168 (Parkhill *et al.*, 2000). The absence of known classical virulence genes, which are usually key for detection and characterisation strategies in other pathogens, poses a further problem for detection of pathogenic campylobacter. However the sequencing of the genome has recognised many putative virulence genes or other candidate gene groups, which could be evaluated for use in detection and characterisation. This gives insights into the unusual characteristics (Park, 2002) of campylobacter compared to other foodborne pathogens, and offers potential for identification of hidden virulence genes (Bereswill & Kist, 2003). Moreover the many Polymerase Chain Reaction (PCR) targets already described could also be evaluated for this purpose (Gonzalez *et al.*, 1997; Lawson *et al.*, 1998; Stucki *et al.*, 1995).

Many genotyping methods have been described for *campylobacter* (Champion *et al.*, 2002; Desai *et al.*, 2001; Duim *et al.*, 2001; Rivoal *et al.*, 1999). These can suffer from a lack of standardisation and variable discrimination, and results can be subjective, leading to misidentification (Wassenaar & Newell, 2000). Some of these shortcomings have recently been overcome by the development of a nucleic acid sequence-based Multi Locus Sequence Typing (MLST) scheme (Dingle *et al.*, 2001).

This has the advantages of a methodology that provides a discriminatory, reproducible molecular profile and provides data, which are directly comparable between laboratories, and standardised nomenclature. Conserved single nucleotide polymorphisms (SNPs) have been identified within the MLST alleles identifying the allelic profile of the major epidemiological lineages and thus providing potential markers for use in detection and characterisation of specific strains.

This investigation proposes to address these issues with the development of strategies for rapid detection of specific campylobacter strains using real time PCR for known molecular targets and SNPs within the MLST alleles. It was intended that this could provide methods applicable for direct detection and characterisation of *campylobacter* within potential sources. Furthermore, recognition of clusters of cases in time and place, which may be due to a common source of infection, would substantially inform epidemiological investigations. Many campylobacter isolates are not identified further than species level. Therefore rapid methods for strain characterisation would be critical to deliver timely epidemiology and enable tracking throughout the food chain.

This section gives a detailed introduction to aspects of campylobacter physiology, disease, epidemiology, population structure, and potential sources of infection, detection and characterisation methods. Additionally a background to the techniques used within this study is described.

1. 2 Background to *Campylobacter* in foodborne disease

Since the first discovery of *campylobacter* in 1972 by Belgium clinical microbiologists (Kirst, 1985) the organism has been isolated in almost every country where investigations have been carried out. *Campylobacters* are regarded as the most common bacterial cause of enteritis worldwide and their prevalence makes them important both from a clinical and economic perspective (On, 1996b).

Despite improvements over the last few years in education with government campaigns and publicity promoting food safety (Attenborough & Matthews, 2000) food poisoning in the UK continues to be an ongoing problem. Over 1 million cases were reported in the year 2000 with an actual incidence close to 10 million (Jones, 2001). Much of this increase is attributed to infections caused by campylobacter. Campylobacteriosis, caused primarily by the species *C. jejuni* or *C. coli* is the most commonly reported bacterial gastrointestinal infection in the UK with 41,406 laboratory-confirmed cases reported in 2003 (www.hpa.org.uk/infections (12/05/04)). However laboratory reporting underestimates the total burden of infection, (Wheeler *et al.*, 1999) with only one in six cases of infectious intestinal disease likely to report to a general practitioner and only a fraction of these reported to national laboratory surveillance. The actual multiplier for cases of campylobacter was estimated as one in eight (Wheeler *et al.*, 1999). Although the disease is usually self-limiting, the high incidence means that campylobacter constitutes a major health issue. Studies of campylobacter epidemiology are hampered by the lack of universally applied typing methods with a standardised nomenclature, which is

possibly one of the reasons for the slow progress in elucidation of transmission routes and epidemiology.

1.3 History and discovery

Many prominent clinical and veterinary scientists contributed towards the discovery of campylobacter, a process which took decades, requiring the development of new diagnostic techniques. The earliest documented description of campylobacter was by Theodor Escherich (1886) who described what appeared to be typical campylobacters in the colons of infants who died of what he called "cholera infantum". Later, in 1906, John McFadyean and Stewart Stockman described what now would be known as *Campylobacter fetus* from a case of epizootic abortion in sheep (McFadyean & Stockman 1913). Plastridge *et al.*, (1947) discovered a similar problem transmitted venerally in carrier bulls, which caused death of the foetus at an early age. Plastridge's work was crucial in control of the infection. His organism was later described by Florent in Belguim in 1959 and is now known as *C. fetus* subspecies *venerealis* (Florent, 1959). The first link between the early campylobacters (then described as vibrios) and diarrhoea was made by Jones *et al* in 1931 in the United States. They managed to obtain pure cultures from cattle with "winter dysentery" and isolate the organism. It was named *V. jejuni* due to it being antigenically distinct from *V. fetus*. The first human campylobacter infections documented (as *V. jejuni*) were, what would now be considered as classic campylobacter outbreaks, including a milk–borne outbreak of diarrhoea which affected 355 inmates of two adjacent state institutions (Levy, 1946).

Elizabeth King in the 1950s made a prominent contribution towards the identification of campylobacters, by making the distinction between the classical *V. fetus* and the strains described by Jones *et al* (1931), which had a higher optimum growth temperature of 42°C. She also recognised the relationship of the organism to diarrhoea, and made the association with a very similar organism in poultry (Dekeyser *et al.,* 1972; King, 1957; King, 1962).

The first isolation of campylobacter from faeces was in 1972, by Dekeyser and Butzler in Belgium (Butzler *et al.,* 1973; Dekeyser *et al.,* 1972). They were able to make the connection between a "related vibrio" from the blood of a previously healthy woman with acute febrile haemorrhagic enteritis and identification of a vibrio from her faeces. As a result of collaborative work between Butzler, Dekeyser and Skirrow in the UK, campylobacters were isolated using filtered faeces onto blood agar. With the later development of selective agars, campylobacter detection in the routine laboratory was achievable (Butzler & Skirrow, 1979; Corry *et al.,* 1995).

1.4 Genus description

The change in name from *Vibrio* to *Campylobacter*, (the name derives from the greek word campylos, meaning curved) by Sebald and Vernon (1963) marked a change in campylobacter taxonomy. The taxonomy of campylobacteria is complex and rapidly evolving, with currently fifteen *Campylobacter* species and six subspecies described (Vandamme & On, 2001) (Table 1.1). The most common species are well defined, although new species have been identified over the last decade, for example, *C. lanienae* isolated from pigs (Logan *et al.,* 2000; Sasaki *et al.,*

2003). The introduction of molecular techniques meant that organisms could be re-classified or moved into another genus such as that of *Helicobacter* or *Arcobacter*. A complete revision of the taxonomy and nomenclature of the genus *Campylobacter* and related organisms was proposed by Vandamme *et al.* (1991). After extensive sequencing of 16S rRNA genes from all known campylobacter species, a sixth rRNA superfamily also known as the epsilon subdivision of the proteobacteria was proposed, containing the family *Campylobacteriacae,* the genus *Helicobacter* and a number of other taxa (Vandamme *et al.,* 1991). Three rRNA homology clusters were identified within this rRNA superfamily the first one constituting the organisms associated with human disease such as *C. fetus, C. lari, C. jejuni,* and *C. coli.* Although these four species tend to be the ones contributing to most cases of enteritis in man, the importance of the other species also associated with gastrointestinal illness may be significantly underdiagnosed as a consequence of inappropriate isolation and identification methods (On 1996a). The techniques described within this study are based entirely around the detection and characterisation of *C. jejuni* and *C. coli,* the two most commonly isolated campylobacters from cases of human infection.

Table 1.1 Table of taxonomic positions and potential sources of each (Vandamme & On 2001),

Taxon **rRNA superfamily VI** **rRNA homology group I**	**Potential source of organism**
C. fetus subsp *fetus* *(Campylobacter fetus* subsp *intestinalis)*	Cattle, sheep
C. fetus subsp *venerealis*	Cattle
C. hyointestinalis subsp *hyointestinalis*	Pigs, cattle, hamsters, deer
C. hyointestinalis subsp *lawsoni*	Pigs
C. consiscus	Humans
C. mucosalis (*C. sputorum* subsp *mucosalis*)	Pigs
C. sputorum biovar *sputorum* (incorporating biovar *bubulus*)	Humans cattle pigs
C. sputorum biovar *faecalis* (*C. faecalis*)	Sheep, bulls
C. sputorum biovar *parahaemolyticus*	
C. curvus (*Wolinella curvus*)	Humans
C. rectus (*Wolinella rectus*)	Humans
C. showae (*Wolinella curva* subsp *intermedius*)	Humans
C. gracilis (*Bacteroides gracilis*)	Human
C. upsaliensis	Dogs cats
C. helveticus	Dogs cats
C. coli	Humans, Pigs, poultry, cattle, sheep, birds
C. jejuni subsp *jejuni*	Humans, poultry, pigs, cattle, sheep, birds, mink, rabbits, insects, dogs.
C. jejuni subsp *doylei*	Humans
C. lari	Birds, river water and seawater, dogs, cats, monkeys, horses.
*C. lanienae**	Pigs

*(Logan *et al.*, 2000)

1.5 Morphology

Campylobacters are curved or "S" shaped slender spiral rods, 0.2-0.8μm wide and 0.5μm to 5μm long, which are gram negative and non-sporeforming (Blaser, 1995). They possess one (monotrichous) or two (amphitrichous) flagella, a polysaccharide capsule (Karlyshev *et al.,* 2001) and demonstrate darting corkscrew-like motility, characteristic of this genus when viewed with phase contrast microscopy. The acute inflammatory enteritis associated with campylobacteriosis commonly extends down the intestine to infect the colon and rectum. For infection to become established campylobacters must be able to first survive the acidic conditions within the stomach and then colonise the ileum and jejunum. The rapid motility and spiral shape of campylobacter and ability to adhere to the mucosal cells make campylobacters highly effective in gut colonisation (Ketley, 1997; Skirrow & Blaser, 2000).

A further physical state of campylobacter has been described, termed the viable non culturable (VNC) state (Cappelier *et al.,* 1997; Tholozan *et al.,* 1999), although much speculation surrounds the validity of this form. It has been described to occur usually in response to starvation (Talibart *et al.,* 2000) or a decline in optimum environmental conditions. The cells undergo a metabolic and morphological change, where they stop being motile and take on a spherical form with apparent loss of culturability (Cappelier *et al.,* 2000). Although non-culturable, it is suggested that even in this state the cells may still present a potent risk of colonisation (Talibart *et al.,* 2000). Studies have also shown that *C. jejuni* can be highly adaptable to changes in environmental conditions, which has resulted in the rapid evolution of some strains as a result of storage, culture, and passage conditions (Gaynor *et al.,* 2004).

1.6 Clinical presentation of infection

Campylobacter infections can often appear with clinical manifestations similar to those seen as a result of salmonella or shigella infection. Campylobacter colonisation of the human intestinal tract usually results in a prodromal period of fever then the onset of diarrhoea, abdominal pain, sickness, fever and malaise within one to five days (Black *et al.,* 1988; Wassenaar & Blaser, 1999). Results from the first year of the *Campylobacter* Sentinel Surveillance Scheme (Gillespie *et al.,* 2001) have shown that 95% of cases presented with diarrhoea, 27% with bloody diarrhoea, 35% had vomiting, 85% had abdominal pain, and 78% fever. The length of illness ranged from a minimum of 1 day to 730 days, with the average being 11 days. Of these cases 10% were hospitalised. These tended to be either the very young or elderly.

Symptomless excreters of campylobacter have been described (Anon, 2000; Butzler, 1982; Porter & Reid, 1980) possibly due to the relative resistance that has been recorded in high-risk groups of people such as those who habitually drink raw milk, or obtain drinking water from a private well. This has also been reported in children without diarrhoea in developing countries, where breast-feeding appeared to have a protective effect (Megraud *et al.,* 1990). Sequential infections within a six to eighteen month time frame have also been documented (Kendall & Tanner, 1982). Experiments with human volunteers have shown that inoculations of 5-500 colony forming units (CFU) were required for high attack rates (Black *et al.,* 1988), but lower doses such as those of waterborne infections were still capable of causing disease. With the low infective dose of campylobacter, it is likely that large waterborne, or milkborne outbreaks are caused by the easy transfer of the bacteria

through the stomach since the milk or water exerts a buffering action, and encourages rapid wash through.

Many associations between campylobacter enteritis and other syndromes or diseases have been made. This includes the reactive arthritis occurring in 1% of cases, leading to syndromes such as Guillain Barré syndrome (GBS) (Kaldor & Speed, 1984), Miller Fisher Syndrome (Yuki, 2001) and Reiter's syndrome (Keat & Rowe, 1991). Associations have been made between campylobacter infection and other conditions, such as immunoproliferative small intestinal disease (Lecuit *et al.,* 2004), pseudoappendicitis, pancreatitis, nephritis and urinary tract infection (Skirrow *et al.,* 1993). Complications can also occur with a campylobacter infection especially when the patient has underlying immune suppression such as in HIV/AIDS. Bacteriaemias have been reported in these cases, and possible meningitis (Skirrow & Blaser, 2000).

GBS, a post infectious immune disease, was first described in 1982 (Kaldor & Speed, 1984), and has been linked to *C. jejuni* infection through controlled epidemiological studies (Tam *et al.,* 2003). It has been associated with other infectious agents including the Epstein Barr virus and Cytomegalovirus (Winer, 2001). Antiganglioside antibodies are induced by *C. jejuni* infection by molecular mimicry between *C. jejuni* lipopolysaccharides (LPS) and gangliosides. Clinically, GBS presents as a rapid progressive neuropathy, caused by the antiganglioside antibodies exerting an effect on the neuromuscular junction (Ang *et al.,* 2002). The campylobacter sialic acid lipopolysaccharide (LPS), which has endotoxic and adherence properties, as well as expressing antigenic variation, is implicated as the potential surface component (Aspinall *et al.,* 1992). The true incidence of cases of *C.*

jejuni mediated GBS are difficult to determine despite controlled epidemiological analyses, mainly due to the absence of the organism in the patient's stools by the time neurological symptoms present. Moreover, the rarity of the condition and lack of standardised typing methods can make cases difficult to compare and study. Particular *C. jejuni* Penner serotypes (Penner & Hennessey, 1980) have been associated with GBS, these being HS:19 globally (Kuroki *et al.,* 1993; Nachamkin *et al.,* 2001) and HS:41 in South Africa (Wassenaar *et al.,* 2000). The *galE* (UDP-galactose 4-epimerase) gene involved in the production of LPS has been identified as a trigger of GBS (Nawaz *et al.,* 2003b), however the progress in understanding campylobacter associated chronic sequelae is far from clear.

1.7 Pathogenesis of infection

Campylobacters are able to enter the intestine *via* the stomach acid barrier and colonise the distal ileum and colon where they initially colonise the mucus and adhere to intestinal cells. Disruption of the normal absorptive processes occurs possibly by damage to the intestinal epithelial cells through the release of toxins, invasion, or by the initiation of an inflammatory response.

(i) Attachment

Attachment to host cells and extracellular matrix proteins is considered a crucial primary event in the pathogenesis of enteritis. Flagellae were originally reported to be the putative adhesions in *C. jejuni* (McSweegan & Walker, 1986). Inactivation studies on the flagellin genes have demonstrated no effect on the adherence to epithelial cells (Grant *et al.,* 1993; Wassenaar *et al.,* 1991). Therefore the flagella genes were proposed to be only important in the preadherence process by the

provision of motility. Additionally LPS has been described to bind to eukaryotic cells (McSweegan & Walker, 1986). Various candidates for adhesion of campylobacter cells within the intestine have been described. The CadF (for campylobacter adhesion to fibronectin) outer membrane protein (37kDa) has been described as being able to promote the binding of *C. jejuni* and *C. coli* to fibronectin (Monteville *et al.*, 2003). Mutants deficient in the *cad*F gene have been shown unable to colonise chickens suggesting that the CadF protein is required for colonisation (Ziprin *et al.*, 1999). Additionally the *peb*1A gene has been demonstrated to play an important role in epithelial cell interactions and has been demonstrated to be located within an operon homologous to those for ABC (ATPase Binding Cassette) transporters in other bacteria (Pei & Blaser, 1993). Moreover mutations within the gene reduce the interactions with epithelial cells and the intestinal colonisation of mice (Pei *et al.*, 1998). Additionally, pilus-like appendages were demonstrated to be produced by *C. jejuni* in response to environmental stimuli, such as a growth environment containing bile salts (Doig *et al.*, 1996). However these were later reported to be a bacteria-independent morphological artifact of the growth medium (Gaynor *et al.*, 2001).

(ii) Iron acquisition

For campylobacters to be able to colonise the intestine, they must be able to survive alongside resident flora and avoid the host immune system (Ketley, 1997). Iron is essential for the growth of most bacteria and the ability to acquire iron under the iron limiting conditions of the digestive tract is seen as a virulence and survival determinant in many bacteria. This is due to the low solubility of iron under physiological conditions, and the limited availability of iron for microbial assimilation due to its sequestration by high affinity iron binding proteins such as

transferrin and lactoferrin (Richardson & Park, 1995). Iron is also vital in the formation of reactive oxygen species enabling pathogens such as campylobacter to maintain intracellular iron homeostasis and cope with oxidative stress. In campylobacter these two functions seem to be under the control of possibly overlapping iron responsive regulatory systems (van Vliet *et al.,* 2002).

The majority of bacteria acquire iron through the release of siderophores which chelate ferric ions, are reinternalised via a transport system, resulting in iron release into the cytoplasm (Salyers & Witt, 1994). Campylobacter do not employ this type of transport system, but compete for siderophores produced by the indigenous microflora of the gut (Baig *et al.,* 1986; Field *et al.,* 1986). It has been proven that one of the methods whereby *C. jejuni* and *C. coli* can acquire iron is from enterochelin, the phenolate-type siderophore produced by *E.coli* (Baig *et al.,* 1986) or from ferrichrome, but not from aerobactin, desferal, ferritin, lactoferrin or transferrin (Pickett *et al.,* 1992).

Several potential iron uptake systems have been described in *C. jejuni*, which have only been recognised recently through comparative sequence and genome analysis of the sequenced *C. jejuni* genome. Systems include; (i) ferrous iron uptake utilising *feo*B (Parkhill *et al.,* 2000), and *cor*A (van Vliet *et al.,* 2002) (ii) Heme uptake (Rock *et al.,* 2001) (iii) Siderophore uptake, including enterochelin and ferrichrome and (iv) other uptake systems including *cfr* A originally identified in *C. coli* (Guerry *et al.,* 1997).

Enterochelin uptake in campylobacter has been described by use of the *ceu* (Cj1352-1355) system, which encodes components of a binding protein dependent cytoplasmic membrane ABC transport system. The *ceuE* component was first described and is thought to be the periplasmic enterochelin binding protein, which acts in conjunction with *ceu*B and *ceu*C (integral membrane proteins) and *ceu*D (ATP binding protein) to allow the successful utilisation of enterochelin (Richardson & Park, 1995).

(iii) Invasion

Both invasive and non-invasive campylobacters have been described (Carvalho *et al.*, 2001). Campylobacter invasion has been strongly implicated by the presence of blood and leukocytes in the stools and many studies have been carried out with experimental cell lines including INT 407 cells (Konkel & Joens, 1989), Caco-2 cells, HEp-2 cells (Everest *et al.*, 1992) and Chinese hamster ovary cells (Klipstein *et al.*, 1985) where invasion *in vitro* is easily demonstrated. Clinical isolates have been described as more invasive than environmental isolates (Konkel & Joens, 1989). Additionally *C. jejuni* invasion has been demonstrated to occur preferentially at the basolateral cell surface (Monteville & Konkel, 2002) when T84 eukaryotic cells were used. The mechanism for invasion of campylobacters is unclear, but there is good evidence that host cell invasion does occur *in vivo*. Although the *fla*A gene itself has been shown not to be directly involved in invasion, functional flagella were described as a requirement for invasion (Wassenaar *et al.*, 1991; Yao *et al.*, 1994). The mechanisms by which *C. jejuni* enter the host cells are also undefined. It is possible that *C. jejuni* may use a microfilament dependent process (De Melo *et al.*, 1989), Konkel & Joens, 1989) as in the case of other invasive pathogens. Others have

reported a microtubule dependent process (Kopecko *et al.,* 2001) or via clathrin coated pit associated receptors and a reaction of the endosome with microtubules (Oelschlaeger *et al.,* 1993). Additionally host clathrin-independent endocytosis and signal transduction involving the mammalian caveolae are implicated as a further mechanism (Wooldridge *et al.,* 1996).

(iv) Toxins

Although toxins have been described for *C. jejuni* the mechanism of action and the importance to the disease process, in conjunction with the other pathogenic mechanisms, remains undetermined. Many cytotoxins in campylobacters have been proposed (Wassenaar 1997). These include a 70-kDa cytotoxin (Goossens *et al.,* 1985; McCardell *et al.,* 1986); a cytotoxin active on Vero/HeLa cells (Florin & Antillon, 1992); a shiga-like toxin (Moore *et al.,* 1988); haemolytic toxins (Hossain *et al.,* 1993; Misawa *et al.,* 1995; Pickett *et al.,* 1992) and a hepatotoxin (Kita *et al.,* 1990; Kita *et al.,* 1992). The existence of all these, and their importance in the pathogenic process, remains to be substantiated. One of the main well-characterised toxins to be described is the cytolethal distending toxin (CDT) (Pickett 2000; Wassenaar & Blaser, 1999). The structural genes for CDT have been cloned and characterised (Pickett *et al.,* 1996). The production of enterotoxins by campylobacter has been described (Florin & Antillon, 1992), yet the reports have never been validated by identification of the genes responsible. It is likely that the reported effects were the effects of other toxins with similar modes of action (Wassenaar, 1997).

(v) Immune response

Campylobacter infection is usually associated with a local acute inflammatory response and involves damage to the intestinal cells. The host inflammatory response may be responsible for mediating the clinical symptoms (Ketley, 1997). A key immune feature of enteric infections is the production of cytokines such as interleukin–8 (IL-8), which is a potent chemotactic factor for immune effector cells. The inflammatory response of *C. jejuni* has been demonstrated to mediate the release of interleukin-8 when exposed to intestinal epithelial cells and also through the release of CDT (Hickey *et al.,* 2000).

Campylobacter disease has been suggested to be an unfortunate outcome of colonisation and is thought to be related to host immune competence (Newell, 2002). This is clear since the organism can act as a commensal in birds, and a transient in most immune humans and animals; only occasionally does colonisation occur and disease ensue. Reduced immunity in humans is proposed to be a result of a lack of exposure, as low-level exposure to campylobacters can result in immune protection. Likewise the exposure to organisms with environmental stresses and therefore impaired invasiveness may have the same effect. The introduction of intensive broiler farm practice has meant the same pathogenic strains are probably being maintained which are able to survive the processing stage (Newell *et al.,* 2001). It is possible that this, combined with the reduction in the time taken from meat processing to the consumer, is unintentionally maintaining the viability of campylobacters, therefore increasing the risk to human health (Newell, 2002).

(vi) Animal models

Problems with the elucidation of the pathogenesis of campylobacter infection are complicated by the absence of suitable animal models. Existing models include the use of ferrets, pigs, calves, primates, cats, dogs, chickens or rats however all are either cumbersome to use, expensive or produce inconsistent results (Hodgson *et al* 1998; Newell, 2001). Success has been achieved with the severe combined immunodeficient mouse (SCID) model, where pathology consistent to that seen within the human intestine has been observed (Hodgson *et al.,* 1998; Young *et al.,* 2000) and the rabbit ileal loop model has been used to demonstrate pathological changes (Everest *et al.,* 1993).

1.8 Campylobacter epidemiology

Campylobacter infection in humans occurs all year around, although seasonal fluctuations arise. The highest infection rates are reported in spring and a further upsurge in cases occurs around autumn in temperate countries (Nylen *et al.,* 2002). In tropical countries there is little seasonal variation except for slightly more infection during the rainy season (Taylor, 1992). Most infections are sporadic in origin, although campylobacter outbreaks occur and nosocomial transmission has been documented (Llovo *et al.,* 2003). Epidemiological studies have established that campylobacters are widespread throughout the food chain, however outbreaks account for only 2% of the total reported cases of infection (Frost *et al.,* 2002) most of which have been from food or water sources. Those traced tend to occur within family clusters or at large social gatherings such as conferences (Raupach & Hundy, 2003), where food hygiene measures have been inadequate resulting in cross contamination.

The highest rates of infection are seen in young children (1-4 years) particularly if residing in a household with puppies, kittens or chickens as pets (Tenkate & Stafford, 2001) and in males aged 25-35 years (Gillespie *et al.*, 2002). Reported incidences showed that the rates of infection were higher in males for all age groups (Frost *et al.*, 2002) and exposure to live farm animals is also hypothesized to increase the risk (Potter *et al.*, 2003).

1.9 Potential sources of human infection

Human infections due to campylobacter primarily occur through the ingestion of undercooked or contaminated foodstuffs. The rare occurrence of outbreaks and sporadic nature of the disease suggests that illness is largely the result of cross-contamination that occurs in the home during food preparation (Brown *et al.*, 1988; De Cesare *et al.*, 2003).

(i) Poultry and other meat products

Poultry has long been associated with campylobacter infection. A survey carried out by the Food Standards Agency reported that 50% of retail poultry in the UK was contaminated* and a case control study estimated that chicken products were responsible for 48% of cases (Harris *et al.*, 1986). There is a short term risk of cross contamination, due to the initial large numbers of campylobacter per carcass and the low infective dose of 5-500 CFU (Black *et al.*, 1988; Robinson & Jones, 1981). It is therefore not surprising that there has been a significant epidemiological link made with the association of illness with undercooked chicken.

*www.foodstandard.gov.uk/ multimedia/pdfs/campsalmsurvey.pdf (24/04/04)

The prevalence of campylobacters in other fresh retail meats has not been so extensively investigated. A small-scale survey of bovine, ovine, porcine liver and chicken portions found 73.2% tested were positive for campylobacter species (Kramer *et al.,* 2000). In this study the prevalence of campylobacter varied according to the meat type with *C. jejuni* predominating in chicken, lamb and ox livers and *C. coli* predominating in pigs liver. Interestingly 30% of all the meat samples investigated yielded more than one strain of campylobacter, determined through subtyping using serotyping (Frost *et al.,* 1998) and phage typing (Frost *et al.,* 1999). The concept of co-infection with more than one type of campylobacter has also been proposed in other studies (Englen & Fedorka-Cray, 2002; Richardson *et al.,* 2002). This emphasizes the critical importance of examining more than one presumptive colony per sample from food products during traditional cultural detection. It is possible that multiple strain campylobacter contamination could lead to misleading results from epidemiological surveys when isolation techniques may promote some strains over others. One of the requirements for this study is that it should be possible to identify specific campylobacter strains from a potentially mixed background of other campylobacters.

(ii) Milk

Milk is another potential source of human campylobacter infection (Kalman *et al.,* 2000; Peterson, 2003). It is assumed that contamination occurs due to faecal contamination during milking, however cases of campylobacter mastitis have also been recorded (Lander & Gill, 1980). Furthermore contamination of milk has been documented from birds (in particular magpies and jackdaws) pecking milk bottle tops and transferring organisms into the milk (Hudson *et al.,* 1991; Riordan *et al.,*

1993; Stuart *et al.,* 1997). Although campylobacter infection from milk is completely preventable through pasteurisation, outbreaks can still occur (Gillespie *et al.,* 2003), especially with those who drink raw milk or where there has been a problem with the pasteurisation process (Peterson, 2003).

(iii) Water

There have been several outbreak reports of outbreak epidemiology implicating water as an important reservoir of campylobacters and potential vehicle of transmission of campylobacter to humans and domestic animals (Clark *et al.,* 2003; Engberg *et al.,* 1998; Obiri-Danso & Jones 1999;Gillespie unpublished 2004). Waterborne outbreaks of campylobacter have been reported from various countries (Bolton *et al.,* 1987) where portable water supplies can be contaminated. Natural waters used for recreational purposes can also be potential risks (Anon, 1998). A study employing molecular techniques for analysis of water samples found campylobacter in water obtained from various locations including swimming pools, tap waters, jacuzzis, and lake waters (Moore *et al.,* 2001). Studies have shown that campylobacters are sensitive to chlorination (Megraud & Serceau, 1990). The presence of campylobacters in tap water or swimming pools must be a result of a breakdown in chlorination practice or faecal contamination from humans in the case of swimming pools (Clark *et al.,* 2003). The ability to survive in drinking water may be strain dependent. Poultry isolates showed prolonged survival compared to isolates from human and water sources (Cools *et al.,* 2003). The contamination of fresh water systems may be due to animal excrement from farm lands entering the water supply or from sewage effluents (Obiri-Danso & Jones, 1999).

(iv) Cross Contamination

Preparation and handling in the commercial and domestic kitchen is a potential risk factor for human infection, with risks of direct hand-mouth exposure and cross contamination of the kitchen environment and other ready to eat foods such as salad products (Kramer *et al.,* 2000). Studies have indicated that relatively high numbers of bacteria may be transferred to a food even 1-2 hours after surface contamination (Moore *et al.,* 2003). Outbreaks of campylobacter occurring in restaurants are unusual, although they do occur. Outbreaks occur as a result of undercooking of food such as in the case of an outbreak involving stir-fried food, where the chicken pieces had been cut too large to facilitate thorough cooking (Evans *et al.,* 1998) or barbecues where food is eaten underdone (Allerberger *et al.,* 2003). The majority of campylobacter infections should be completely preventable by proper cooking, handling and storage of high-risk foods such as poultry and raw meats.

Lack of education is a key factor in the increase in the cases of campylobacter infection. It was reported from Iceland that, after the introduction of fresh chicken into the diet, campylobacter cases increased dramatically, due to the general population of Iceland not being aware of how to handle or cook fresh chicken (Reierson *et al.,* 2001). Barbecues appear to present hazards for infection, because they permit easy transfer of bacteria from raw chicken to fingers and other foods and from these to the mouth (Butzler & Oosterom, 1991). They also represent a significant cause of campylobacter infection due to either the undercooking of chicken meat in the centre when the outside appears cooked or through cross contamination from raw chicken to ready to eat food items such as salads, or utensils. It has been suggested that poultry, compared to other meats, is handled differently in

the kitchen. Up to 56% of poultry packaging is contaminated with campylobacter, suggesting that cross-contamination may occur between products, which are bought at the same time or stored together (Jorgensen *et al.*, 2002).

Reports from the *Campylobacter* Sentinel Surveillance Scheme (Gillespie *et al.*, 2001) linking patient questionnaires and the foods eaten in the two weeks prior to getting ill, compared with the national food survey, suggested that chicken was not the highest risk factor for infection per person given the volume consumed. In this analysis the highest risk was associated with eating paté, followed by pre-packed sandwiches and meat pies. Campylobacters have become successful bacteria by evolving to colonise many hosts and adapting to survive in different environmental conditions, such as many food products.

1.10 Reservoirs for campylobacter

Campylobacters appear to have evolved for optimal growth within the avian gut as a commensal, although they are also harmless commensals within the gastrointestinal tract of many domestic animals, food-producing animals such as cows, sheep (Ertas *et al.*, 2003; Stanley & Jones, 2003), goats (Hutchinson *et al.*, 1985) and pigs (Moore & Murphy, 2003) and other animals including reindeer (Hanninen *et al.*, 2002). They have also been isolated from a wide range of environmental locations including the sand of bathing beaches (Bolton *et al.*, 1999), soil (Santamaria & Toranzos, 2003) and water, probably due to contamination from animal faeces (Bolton *et al.*, 1987).

Most farm animals are not colonised at birth but soon become colonised by contact with wild animals or birds, or by the transfer from older members of the herd or flock

(Stanley & Jones, 2003). The relatively few strains circulating within a herd at any one time suggests that transmission between animals is a relatively common process and may influence the seasonality of human campylobacter infection (Stanley & Jones, 2003). The numbers of campylobacters carried within the intestines of lambs was described to vary throughout the year with a peak in May and trough in October (Stanley *et al.,* 1998a; Stanley *et al.,* 1998b). This was reflected in the risk of carcass contamination at slaughter, and transmission of the campylobacter through the food chain.

Campylobacter is thought to colonise chickens during rearing (horizontally), although transmission of campylobacter through the egg (vertically) has been described as a very rare event (Corry *et al.,* 2001; Sahin *et al.,* 2003). Newly hatched chicks are uncolonised, but asymptomatic infection is detectable in chickens at 2-3 weeks of age (Evans & Sayers, 2000; Wallace *et al.,* 1998). Rapid spread of infection can occur throughout the flock within 3 days and transmission via water, both within the environment and broiler houses (Clark *et al.,* 2003; Evans *et al.,* 2003; Newell & Fearnley, 2003; Said *et al.,* 2003) as well as contact with farm workers who work with other animals (Altekruse & Tollefson, 2003) are described as a potential sources of infection. Zimmer *et al* (2003) described waterborne campylobacter strains from a commercial broiler house to be different to those found in the chickens suggesting water is not the key contributor of infection. Also, there appears to be a seasonal prevalence of campylobacter in flocks, where some flocks remain campylobacter infection free during the cooler months of the year (Freidman *et al.,* 2000; Willis & Murray, 1997). Associations have been made with particular types of air conditioning systems within broiler houses and the infestation of insects, especially

beetles (Refregier-Petton *et al.,* 2001). Furthermore, birds including sparrows (Chuma *et al.,* 2000) and possibly rodents (Gregory *et al.,* 1997) have been described to be a potential source of contamination. Careful attention to biosecurity within the broiler house such as restricting access, changing clothes, and eliminating vermin, has been shown to be important in the control (Gibbens *et al.,* 2001).

1.11 Growth and persistence in the environment

Campylobacters in the laboratory appear to be highly fragile organisms and very susceptible to environmental stresses, especially when compared to other food borne organisms (Park, 2002). Campylobacters are microaerophilic; this delayed their recognition as pathogens and probably hampers accurate measurements of their true incidence today (Hu & Kopecko, 2002). They grow optimally in atmospheres containing 5% oxygen and a temperature range of 37-42°C. However, the organism has been reported to be able to adapt to an aerobic metabolism when required by the environment (Jones *et al.,* 1993). This has been confirmed through genome studies where a complete citric acid cycle and a complex and highly branched respiratory chain have been described, which allows for both aerobic and anaerobic respiration (Kelly, 2001). These properties enable the organisms to survive in a range of environments including the avian and mammalian gut and food products (Manfreda *et al.,* 2003b). Although campylobacters do not grow outside the host they can remain viable for long periods in water, foods and within environmental samples (Newell, 2002).

The persistence of campylobacter throughout the domestic kitchen, on fingertips and within other environments has been described (Bolton *et al.,* 1999; Coates *et al.,*

1987; Humphrey, 2001; Humphrey *et al.,* 2001). However campylobacters are very sensitive to drying and consequently do not survive well on dry surfaces (Mattick *et al.,* 2003) Fermandez *et al.,* 1985). Osmotic stress, pH and high temperatures pose more of a problem to campylobacters than to other food borne organisms such as *Salmonella typhimurium* and *Listeria monocytogenes*. Campylobacters will not grow in concentrations of sodium chloride of 2% (Doyle & Roman, 1982), are incapable of growth below pH 4.9 (Blaser *et al.,* 1980), and will not survive food preparation such as adequate cooking by maintaining a core temperature of 70°C for 2 minutes or more, (Attenborough & Matthews, 2000) or pasteurisation (Robinson & Jones, 1981).

1.12 Campylobacter detection methods

Campylobacter detection has traditionally depended on the isolation of bacteria with selective media followed by identification and determination of antimicrobial susceptibility patterns (Winters *et al.,* 1997). Various campylobacter selective media have been described, including the original Skirrow medium (Skirrow, 1977) used for the initial isolation, Preston medium (Bolton & Robertson, 1982), cefoperazone amphotericin teicoplanin (CAT) medium, cefalozin charcoal deoxycholate (CCD) medium (Bolton *et al.,* 1984b) and the commonly used modified CCD medium including cefoperazone in place of cefalozin (Hutchinson & Bolton, 1984). Most media have been developed to allow incubation at 37°C and include ingredients (e.g. defibrinated blood, or charcoal) designed to protect campylobacters from the toxic effect of oxygen derivatives (Corry *et al.,* 1995). These have permitted the isolation of campylobacter from a matrix of other bacteria, and allowed for the growth of campylobacter without the overgrowth of other more common bacteria.

Many methods have been described for the separation of bacteria from food matrices, by chemical, physical and biological methods with varying degrees of success (Benoit & Donahue, 2003). Common methods aiming to concentrate the bacterial cells include centrifugation either at slow speed to sediment the food particles, leaving the bacteria in solution, or alternatively at higher speeds to sediment the bacteria, although this can often also sediment particulate matter. Filtration has also been used for concentrating bacteria, immobilisation strategies using various substrates (Lucore *et al.,* 2000) and immunomagnetic separation technologies (Liu & Li 2002; Reinders *et al.,* 2002). With all these methods good bacterial cell recovery is difficult to achieve and enrichment culture is often still required.

Detection of campylobacter from products where the campylobacter has been present in low numbers (e.g. food and water samples) has traditionally relied upon a prior selective enrichment step in a liquid enrichment broth such as Bolton or Preston (Bolton & Robertson, 1982) to increase the numbers of organisms for adequate growth on media (Bolton *et al.,* 1986; Hutchinson & Bolton, 1983). Current enrichment and cultural methods for detecting campylobacters from complex matrices such as food, clinical or environmental samples are slow and tedious and often not a true representation of the contamination within the sample due to the presence of cells too damaged or distressed to grow in artificial conditions. Additionally selective media can be inhibitory to certain species preventing growth. This combined with the relative chemical inertness, the slow growth of campylobacters and their inability to ferment or oxidise carbohydrate substrates means that the normal biochemical tests used for the identification of bacteria are

limited, therefore newer comprehensive detection methods are to a great extent overdue.

Many of the molecular methods described for identification and typing of campylobacter do not necessitate cultured cells on a plate, only the genomic DNA is required. Of great utility in the detection of specific campylobacters would be the detection of the presence of specific campylobacter genomic DNA in a suspect sample rather than waiting for bacterial growth in culture to indicate campylobacter contamination. Direct detection of specific DNA would present a more accurate picture of the presence of any campylobacter within a sample. It would not be subject to the problems of a lack of bacterial growth due to the cells being damaged or stressed but the organism would still be detectable by the presence of their DNA. Some methodologies for direct DNA extraction from what once were considered difficult matrices such as faecal samples, food samples and water have been described (Nogva *et al.,* 2000b) which enable the direct isolation of DNA from the sample itself, or the enrichment broth without first culturing the material. It is potentially possible to use a DNA isolation method directly from a suspect starting material in conjunction with a campylobacter specific molecular technique (described in the next section) to identify the campylobacter DNA and therefore confirm the presence of campylobacter. Additionally, strain characterisation could be achieved in the same single process, depending on the capacity of the molecular identification method used.

1.13 Campylobacter DNA isolation for molecular techniques

The use of appropriate DNA extraction methods or kits is vital for successful and accurate PCR or molecular studies on food and faecal samples. Methods have to be chosen carefully in relation to the sample type tested (McOrist *et al.*, 2002). Methods for DNA extraction have progressed from the established isolation methods such as the hexadecyltrimethyl ammonium bromide (CTAB) method to DNA extraction kits, including completely automated robotic extraction systems. Convenient kits whereby DNA of a high purity and quantity can be obtained quickly from culture are commonplace, with kits specifically designed for other sample starting materials such as food and faecal samples also available.

Any food matrix often contains large populations of indigenous microflora, which, although they pose no significant risk to health, their physical presence may mask the presence of campylobacter, which is likely to be present in very low numbers. PCR based methods would seem to be applicable due to greater sensitivity (Englen *et al.*, 2003). Successful application of the PCR assays to food samples has been hindered by the lack of convenient and relatively simple methods for direct detection of PCR products as well as by the lack of rapid and efficient methods for the preparation of PCR-amplifiable DNA from foods (Chen *et al.*, 1997). Much of the difficulty in implementing PCR for the analysis of food samples lies in the problems encountered during the preparation of template DNAs from food matrices. The DNA has to be extracted in sufficient quantities and it has to be pure, the inhibitory substances present in enrichment media and food itself can adversely effect the reaction, lowering the detection capacity. Basically the more complex a food is in terms of its content the more difficult it becomes to detect the bacteria by molecular methods.

Other problems include the technical complexity of PCR technology and insufficient data on its performance in the analysis of naturally contaminated foods (Chen *et al.*, 1997). More recently, kits based on magnetic glass (bacteria binding) bead technologies have been developed and evaluated for the use in isolating bacterial PCR amplifiable DNA directly from the food sample (Nogva *et al.*, 2000b).

Water samples pose a different problem since here the organism is present in small quantities within large volumes of water; therefore enrichment culture would usually be employed. Extended analyses of water by conventional culturing techniques are often unsuccessful in isolating campylobacters (Moore *et al.*, 2001) where results are highly dependent upon the types of isolation media used (Cools *et al.*, 2003). Failure to detect implicated organisms in water may be due to the cells being damaged, due to the time taken from collection to testing or due to the low numbers present in large volumes of water. For these reasons more accurate water testing by isolation of campylobacter DNA and PCR methodologies would give a clearer representation of the true incidence of campylobacter contamination.

1.14 Campylobacter identification methods

C. jejuni and *C. coli* are the most likely campylobacter species to be implicated in human infection therefore the majority of identification methods and typing schemes have been developed for one or both of these species. The differentiation between *C. jejuni* and *C. coli* is interesting since the two species show a high degree of relatedness in protein profiles (Owen *et al.*, 1988) and DNA homology (Roop *et al.*, 1984) and they are often difficult to distinguish due to the limited number of tests available. Moreover, data has suggested that phenotypic and biochemical assays are

unreliable for speciating campylobacters particularly those recovered from non-human sources (Burnett *et al.*, 2002; Engvall *et al.*, 2002).

Human infections by other campylobacter species have been documented, although these are infrequent (Hudson *et al.*, 1991) and are often seen in patients with other conditions (Krause *et al.*, 2002) or those immunocompromised (Woo *et al.*, 2002). Phenotypic identification of campylobacter has depended upon biochemical tests such as oxidase, catalase and biotyping schemes (Bolton *et al.*, 1984a) and differentiation of *C. jejuni* and *C. coli* has traditionally depended on the hippurate hydrolysis test. The hippurate test is commonly used to differentiate between *C. jejuni* and *C. coli,* based upon the presence of N-benzoylglycine amidohydrolase (hippuricase) in *C. jejuni,* but not in *C. coli.* The *hip*O gene encodes the enzyme hippuricase (Hani & Chan, 1995), which cleaves N-benzoglycine (hippurate) into benzoic acid and glycine and can be detected by use of a ninhydrin, based system. However hippurate negative strains of *C. jejuni* have been described (Totten *et al.*, 1987; Thwaites *et al* unpublished), which may have significant implications for the use of this test in epidemiological investigations.

Many other methods have been described for use in identifying campylobacters including serological tests (e.g. Campyslide, BBL Microbiology Systems, Microscreen, Mercia Diagnostics) lectin agglutination (O'Sullivan *et al.*, 1990), cellular fatty acid profiling (Lambert *et al.*, 1987), mass spectrometry (Winkler *et al.*, 1999), PCR techniques and nucleic acid probes (Stonnet & Guesdon 1993; Van Camp *et al.*, 1993). However, as the taxonomy of the genus is rapidly evolving, no

identification schemes have been developed which are applicable for all described species and standardisation between laboratories is difficult to achieve.

1.15 The use of PCR for campylobacter identification

The difficulties surrounding campylobacter detection and identification makes the PCR approach a more favourable option, with the added advantages of increased sensitivity, low cost, ease of use and increased validity of results. Additionally the organism is usually not found in healthy humans therefore PCR identification of campylobacter specific gene regions could provide reliable confirmation of infection. Of utmost importance in the designing of PCR assays is the availability of accurate sequence data. The problem of inaccurate sequence data has been highlighted in the accumulation of chimeric 16S rRNA sequences within public databases (Hugenholtzt & Huber, 2003). The sequencing of the *C. jejuni* genome (Strain NCTC 11168) (Parkhill *et al.,* 2000) has permitted the development of many more PCR assays, which are based on regions of genes specific to a species. For a species where the whole genome sequences are not available (e.g. *C. coli* or *C. fetus*) PCR assays have been based on particular genes (or divergent parts of genes), which have been identified as unique to the species by comparison with the same gene from *C. jejuni* or other species.

Over the last decade a multitude of PCR assays have been described for the identification of *Campylobacter* species and related organisms (Table 1.2). The use of PCR to determine species is the most common and accurate method for speciation of campylobacters and many different targets specific to different species have been described (Table 1.2). This includes the detection of 16S and 23S rRNA genes

although data has suggested that there is insufficient sequence variation in the 16S rRNA gene to discriminate this closely related group of bacteria (Burnett *et al.*, 2002). Likewise the hippuricase (*hip*O) gene can be detected by PCR to confirm *C. jejuni* however the occurrence of atypical hippurate negative *C. jejuni* strains leads to inaccuracies (Totten *et al.*, 1987; Thwaites *et al* unpublished). Uniplex or multiplex assays have been described whereby *C. coli, C. jejuni* and other species can be identified often in conjunction with each other or other related organisms such as *Arcobacters* (Al Rashid *et al.*, 2000). The majority of PCR assays are based around the presence of specific regions of genes known to be unique to the species due to the lack of known virulence factors for campylobacters. Regions of the 16S rRNA gene are commonly described, which has been used both in a uniplex assay for the identification of *C. lari* (Oyarzabal *et al.*, 1997b), and in a multiplex assay for confirmation of the genus (Denis *et al.*, 1999). For *C. jejuni* the *map*A and *hip*O genes are described extensively and for *C. coli* the *ceu*E or aspartokinase gene have been well described.

PCR methodologies have also been described which involve a specific PCR then subsequent methods such as the use of species specific probes (Van Doorn *et al.*, 1997), or a Restriction Fragment Length Polymorphism (RFLP) strategy with a combination of restriction enzymes for distinguishing the species (Cardarelli-Leite *et al.*, 1996; Hurtado & Owen 1997; Jackson *et al.*, 1996). However, PCR RFLP and hybridisation strategies tend to be tedious and lengthy, and also have questionable reproducibility when applied with different reagents, and running conditions in different laboratories making standardisation difficult. Additionally microarray based identification of *C. jejuni, C. coli, C. lari* and *C. upsaliensis* has been described

(Volokhov *et al.*, 2003). This used 5 specific regions in the target genes (*fur*, *gly*A, *cdt*ABC, *ceu*B-C, *fli*Y) and provided a rapid and accurate differentiation assay for the simultaneous differentiation of these species.

Specific conventional PCRs, which aim to distinguish the species based on the occurrence/non-occurrence of amplification products, are more reproducible and quicker to perform, without the reliance on the time-consuming restriction analysis. They also have greater potential for standardisation between laboratories by the incorporation of standardised reagents, conditions and an electrophoresis stage. All the various PCRs described are difficult to evaluate, as they have all been developed and utilised for specific objectives and only some have been extended to show identification from a non-culture sample.

Table 1.2 Common PCRs described for the identification of campylobacters

Gene/combination of genes used	Reference	Type of assay	Genus /species distinguished
Random	(Gisendorf *et al.*, 1992)	multiplex	*C. jejuni and C. coli*
23S rRNA	(Eyers *et al.*, 1993)	multiplex	*C. jejuni, C. coli, C. lari, C. upsaliensis*
Random	(Stonnet & Guesdon, 1993)	uniplex	*C. jejuni*
Intragenomic region of *flaA* and *flaB*	(Wegmuller *et al.*, 1993)	multiplex	*C. jejuni and C. coli,*
MapA	(Stucki *et al.*, 1995)	uniplex	*C. jejuni*
Random	(Winters & Slavik, 1995) (Winters *et al.*, 1997)	uniplex/nested	*C. jejuni*
Random	(Day *et al.*, 1997)	uniplex	*C. jejuni*
CeuE	(Gonzalez *et al.*, 1997)	multiplex	*C. jejuni* and *C. coli*
flaA	(Harmon *et al.*, 1997)	multiplex	*C. jejuni and C. coli*
16S rRNA	(Lawson *et al.*, 1997)	multiplex	*C. upsaliensis* and *C. helveticus*

Table 1.2 (cont) Common PCRs described for the identification of campylobacters

16S rRNA, *hip*O, aspartokinase gene	(Linton *et al.*, 1997)	multiplex	*C. jejuni and C. coli,*
Random fragment	(Ng *et al.*, 1997)	uniplex	*C. jejuni*
16S rRNA	(Oyarzabal *et al.*, 1997b)	uniplex	*C. lari*
Random DNA fragments	(Vandamme *et al.*, 1997)	multiplex	*C. jejuni and C. coli*
sapB2	(Casademont *et al.*, 1998)	uniplex	*C. fetus*
16S rRNA *mapA, ceuE*	(Denis *et al.*, 1999)	multiplex	Campylobacter genus band, *C. jejuni* and *C. coli*
CadF	(Konkel *et al.*, 1999)	multiplex	*C. jejuni* and *C. coli*
*Gly*A	(Al Rashid *et al.*, 2000)	multiplex (hybridisation)	*C. coli, C. lari, C. upsaliensis, Arcobacter butzleri*
16S rRNA	(Thunberg *et al.*, 2000)	uniplex	*C. jejuni*
16S rRNA and *hip*O	(Bang *et al.*, 2002)	multiplex nested PCR	*C. jejuni* and *C. coli*
*hip*O, 23S rRNA, *gly*A, *sap*B2	(Wang *et al.*, 2002)	multiplex	*C. jejuni C. coli, C. lari, C. upsaliensis C. fetus*

1.16 The use of PCR for identification of campylobacter from other matrices

It is recognised that the time taken for analysis of food samples for campylobacter by conventional microbiology techniques is too long, hence PCR techniques have been employed. Despite this, the techniques required to prepare samples for PCR testing are still time consuming and usually involve culture, therefore any damaged or stressed cells will still be missed. Currently the most successful methods for isolation of campylobacter from food samples involve enrichment techniques, culture, DNA extraction and subsequent PCR (Table 1.3). Many of the PCR assays for campylobacter identification, (Table 1.2) have been applied to investigate food samples (Table 1.3) and an additional number have been solely described for the PCR testing of food samples (Gisendorf *et al.,* 1992; Wegmuller *et al.,* 1993). In most cases a combination of enrichment methods and conventional culturing has had to be employed before the PCR assay is carried out. Although descriptions of direct DNA extraction from the food sample itself (Wegmuller *et al.,* 1993) have been described. Different DNA extraction methods have been employed these ranging from simple procedures involving centrifugation and Triton X-100 treatment (Winters *et al.,* 1997; Winters *et al.,* 1998) to more traditional time consuming methods such as the CTAB (Wilson *et al.,* 1987b) method used by Magistrado *et al* (2001). Techniques involving buoyant density centrifugation have also been employed (Thunberg *et al.,* 2000; Wang *et al.,* 1999) and the use of hydrophobic grid membrane filters (HGMFs)(Wang, 2002).

Despite vast numbers of assays describing the development of PCR assays for campylobacter and the various extraction methods, which have been described, very few have been commercialised into kit form for use on food samples. This is

explained by the fact that PCR assays are very subject to interference by food matrices and components in foods which inhibit PCR reactions, as a result most sample types still require a cultural step to enhance assay performance and distinguish viable from non viable cells (Feng, 1997). The Bax® system (Manfreda *et al.,* 2003a) is the only described fully automated system whereby a food sample can be loaded into the machine with tablet of PCR reagents, automated PCR is carried out and a result is automatically produced as to the presence or absence of *C. jejuni* or *C. coli.* However this system has certain drawbacks in that only the Bax® system primers can be used, and the machine takes approx 12 hours to run.

Table 1.3 Common PCRs used for detection of Campylobacters and descriptions of the application to food or environmental samples

PCR Assay used (described by)	Methods used in conjunction with PCR	Genus /species distinguished	Type of samples tested	Reference(s)
16S rRNA (Gisendorf *et al.*, 1992)	Enrichment and culture Enrichment and DNA extraction	*Campylobacter* species.	Chicken products	(Gisendorf *et al.*, 1992)
	Culture DNA Extraction (CTAB, Sodium acetate)	*C. jejuni and C. coli*	Duck and chicken	(Magistrado *et al.*, 2001)
Flagellin genes (Nachamkin *et al.*, 1993)	Direct DNA Extractor™ Columns	*C. jejuni*	Human faeces	(Waegel & Nachamkin, 1996)
Intragenomic region of *fla*A and *flab* (Wegmuller *et al.*, 1993)	DNA extraction from foods (Allmann *et al.*, 1995)	*C. jejuni and C. coli*	Milk cheese and yoghurt	(Wegmuller *et al.*, 1993)
	Enrichment/non enrichment Immuno magnetic separation or Sedimentation	*C. jejuni and C. coli*	Environmental water, sewage, food samples (minced beef chicken and pork)	(Waage *et al.*, 1999)
Random (Winters & Slavik, 1995)	Culture centrifugation and Triton X-100 treatment	*C. jejuni*	Chicken washes	(Winters & Slavik, 1995)
	Culture centrifugation and Triton X-100 treatment, nested PCR	*C. jejuni*	Vegetables and fruit	(Winters *et al.*, 1998)

Table 1.3 (cont) Common PCRs used for detection of Campylobacters and descriptions of the application to food or environmental samples

16S rRNA (Linton *et al.*, 1996) IVS in 16S rRNA (Linton *et al.*, 1994)	Membrane Filtration, selective culture DNA extraction by Boom method (Boom *et al* 1990)	*C. upsaliensis* and *C. helveticus*	Seeded faecal samples	(Lawson *et al.*, 1997)
16S rRNA, *hip*O, aspartokinase gene (Linton *et al.*, 1996)	Direct from faeces by Boom (Boom *et al.*, 1990)	*C. jejuni* and *C. coli*	Human faeces	(Lawson *et al.*, 1998)
*ceu*E (Gonzalez *et al.*, 1997)	Enrichment, culture DNA extraction by traditional Proteinase K and phenol	*C. jejuni* and *C. coli*	Gall bladders of sheep and cattle	(Ertas *et al.*, 2003)
16S rRNA *hip*O gene Aspartokinase gene	Culture and DNA Extraction (CTAB)	*C. jejuni* and *C. coli*	Human faeces	(Linton *et al.*, 1997)

Table 1.3 (cont) Common PCRs used for detection of Campylobacters and descriptions of the application to food or environmental samples

Random fragment (Ng *et al.*, 1997)	Hybridisation PCR	*C. jejuni*	Chicken rinse	(Ng *et al.*, 1997)
	Enrichment, Buoyant density centrifugation	*C. jejuni*	Chicken rinse	(Wang et al., 1999)
	Culture DNA Extraction (CTAB, Sodium acetate)	*C. jejuni*	Duck and chicken	(Magistrado *et al.*, 2001)
	HGMF PCR	*C. jejuni*	Chicken	(Wang 2002)
Random (Winters *et al.*, 1997)	Culture centrifugation and Triton X-100 treatment nested PCR	*C. jejuni*	Chicken rinses	(Winters *et al.*, 1997)
	Culture centrifugation and Triton X-100 treatment nested PCR	*C. jejuni*	Vegetables and fruit	(Winters & Slavik, 2000)

Table 1.3 (cont) Common PCRs used for detection of Campylobacters and descriptions of the application to food or environmental samples

Random (Stonnet & Guesdon 1993) Random (Winters *et al.*, 1997) *Hip*O, aspartokinase (Linton *et al.*, 1997) 16S rRNA (Oyarzabal *et al.*, 1997a)	Enrichment Culture Instagene DNA isolation	*C. jejuni, C. coli* and *C. lari*	Sparrow faeces	(Chuma *et al.*, 2000)
Random (Winters & Slavik 1995) Aspartokinase (Linton *et al.*, 1997)	Filter-enrichment culture	*C. jejuni* and *C. coli*	Human faeces	(Misawa *et al.*, 2002)
16S rRNA, *ceuE, mapA* (Denis *et al.*, 1999)	Enrichment, culture, DNA extraction (QIAmp Tissue Kit)	*C. jejuni and C. coli*	Chicken-legs, breast, gizzards, wings, escalopes	(Denis *et al.*, 2001)
	Enrichment, DNA extraction(Jackson *et al.*, 1996)	*C. jejuni and C. coli*	Environmental samples (pig and poultry swabs)	(Manfreda *et al.*, 2003a)
16S rRNA (Vanniasinkam *et al.*, 1999)	Enrichment, DNA Extraction (Saunders *et al.*, 1990)	*C. jejuni*	Clinical specimens -faeces	(Vanniasinkam *et al.*, 1999)

Table 1.3 (cont) Common PCRs used for detection of Campylobacters and descriptions of the application to food or environmental samples

16S rRNA (Thunberg *et al.*, 2000)	Enrichment, culture DNA extraction (Percoll method(Lindqvist *et al.*, 1997), TE buffer dilutions)	*Campylobacter* species.	Broccoli, crabmeat, mushroom, oyster	(Thunberg *et al.*, 2000)
(Winters & Slavik, 2000)	Culture centrifugation and Triton X-100 treatment nested PCR	*C. jejuni and A.butzleri*	Ready to eat foods, dairy products	(Winters & Slavik, 2000)
(Sails *et al.*, 2001)	Enrichment, Prepman DNA extraction	*C. jejuni* and *C. coli*	Environmental waters	(Sails *et al.*, 2002)
(Bolton *et al.*, 2002)	Enrichment culture, PCR ELISA Technique	*C. jejuni* and *C. coli*	Food (meat, shellfish, milk)	(Bolton *et al.*, 2002)
*map*A for *C. jejuni*, *ceu*E for *C. coli*. Regions of 16S rRNA, and 23S rRNA for identification of other species.	Direct from faeces QIAamp mini DNA stool minikit	*C. jejuni, C. coli, C. fetus, C. hyointestinalis.*	Bovine faeces	(Inglis & Kalischuk, 2003)
16S/23S rRNA PCR with DNA probe membrane based colorimetric assay.	Direct from faeces QIAamp mini DNA stool minikit	*C. jejuni, C. coli* and other *Campylobacter* species	Human faeces	(Maher *et al.*, 2003)
Bax® system primers	Enrichment Bax® system (Qualicon)	*C. jejuni and C. coli*	Environmental samples (pig and poultry swabs)	(Manfreda *et al.*, 2003a)

1.17 Phenotypic methods for subtyping

The use of sub typing is an essential approach to identifying specific campylobacter strains and for elucidating the nature and diversity of strains within potential sources and determining their epidemiological significance. To successfully detect from suspect sources, it is not only necessary to identify the presence of *Campylobacter* species but the detection has to go a step further to be able identify a specific strain. This is a particular consideration when the strain of interest may be amongst a diverse background of other types. This allows for the recognition of the link between a case of infection and suspect source of a specific campylobacter.

1.17.1 Serotyping and phagetyping

Serotyping methods have largely been used for epidemiological typing of campylobacter and only more recently have molecular techniques been applied. Two methods for serotyping isolates have been used extensively, the Penner scheme based on heat stable antigens (Moran & Penner, 1999) and the Lior scheme based upon heat labile antigens (Lior, 1982). Serotyping in the Campylobacter Reference Unit (CRU), is based upon an adaptation of the Penner technique and involves the identification of heat stable (HS) antigens by absorbed antisera utilising direct whole cell agglutination (Frost *et al.,* 1998). This scheme defines 48 heat stable serotypes in *C. jejuni* and 17 in *C. coli.* Levels of typeability (70%) are of concern and there is a further requirement to subtype within the most common serotypes. Elucidation of the molecular basis of campylobacter serotyping is an area of great interest, which would provide an advantageous alternative to current serotyping methodology. However progress has been slow for campylobacter, especially in comparison to other organisms, such as in the case of salmonella where large amounts of sequence

information have been obtained for the H and O antigens (Anon 2004; McQuiston *et al.*, 2004).

Phage typing in the reference laboratory is used as an adjunct to serotyping to increase discrimination (Frost *et al.*, 1999). Phage typing schemes have been described by Grajewski *et al* in the USA (Grajewski *et al.*, 1985), Salama *et al* in the UK (Salama *et al.*, 1990) and Khakhira and Lior in Canada (Khakhira & Lior 1992). All three schemes share phages, although the UK scheme is the one used as an adjunct to serotyping in the reference laboratory (Frost *et al.*, 1999). A total of 76 defined phage types have been described. A phage type is described as two or more epidemiologically unrelated isolates giving the same phage reaction pattern. For this scheme reference isolates and propagating strains are available from the National Collection of Type Cultures. Serotyping with the LEP scheme classifies strains into broad groups and 53% of isolates are accounted for by the main 10 serotypes. With the addition of phage typing these groups can be broken down further into between 6 and 29 subtypes, however within some serotypes certain phage types predominate possibly indicating a close clonal relationship within the groups (Newell *et al.*, 2000).

Sero and phage typing combined give an adequate level of discrimination and a typeability of 97% (Frost *et al.*, 1999). These techniques have been in use routinely for phenotypic subtyping. However the usefulness of both techniques is limited by the occurrence of non-typeable organisms and cross reactivity. Additionally, the production and maintenance of the panel of reagents required is time consuming, costly and impractical for use in most clinical laboratories. The success of Kauffman-

White scheme (Popoff *et al.*, 2001) for serotyping of *Salmonella enterica* isolates was attributed to the clonal population structure of the organism, where variation in the immunological reactivity of surface polysaccharides proved a convenient way of establishing clones. However *C. jejuni* and *C. coli* exhibit weakly clonal populations and engage in frequent lateral gene transfer (Dingle *et al.*, 2002), especially the loci encoding characteristics such as flagella and capsular antigens, which are under diversifying selection. This results in horizontal genetic exchange in organisms that do not necessarily share a common ancestor and explains some of the problems in understanding the pattern of campylobacter disease in humans.

1.18 Molecular methods for subtyping

Many techniques have been described for the molecular typing of campylobacter including Pulsed Field Gel Electrophoresis (PFGE) (Anon 2001; Olsen *et al.*, 2001; Ribot *et al.*, 2001), Amplified Fragment Length Polymorphism Techniques (AFLP) (Duim *et al.*, 1999, 2001; On & Harrington 2000), Random Amplified Polymorphic DNA (RAPD) Analysis (Ono *et al.*, 2003) and various PCR Restriction Fragment Length Polymorphism (RFLP) techniques based on different genes and using a variety of restriction enzymes (Shi *et al.*, 2002; Smith *et al.*, 2000a). Most commonly used are the techniques of Pulsed Field Gel Electrophoresis (PFGE) and various formats of Amplified Fragment Length Polymorphism techniques (AFLP) using a combination of one or two restriction enzymes, and either fluorescently labelled sequence based detection of fragments (FAFLP) (Desai *et al.*, 2001) or agarose gel based fragment detection (Champion *et al.*, 2002).

1.18.1 Pulsed Field Gel Electrophoresis (PFGE)

PFGE is a modification of the more traditional Restriction Fragment Length Polymorphism (RFLP) method, which can be used for further strain discrimination when required. Conventional agarose gels are only capable of separating nucleic acids up to a size of 20kb. This is due to the large DNA fragments being unable to migrate through the gel matrix by the same sieving mechanism that separates smaller fragments. The only way that large fragments can move through the agarose gel is by "squirming" through the gel in a parallel direction to the electric field. The technique PFGE takes advantage of this property of DNA and is capable of resolving fragments of 100-1000kb by intermittently changing the direction of the electric field. This is done in such a way that the large DNA fragments re-align more slowly with the new field direction, than the smaller molecules. The electrodes are placed around the perimeter in a hexagonal arrangement, producing alternating currents of 120 degrees and a contoured clamp homogenous electric field (CHEF) (Miesfield, 1999). DNA for PFGE cannot be used in solution but has to be immobilised in agarose blocks, to prevent shearing. These DNA containing blocks are soaked in lysis buffer containing detergent and proteinase K, which purifies the DNA without exerting too much physical force from pipetting. The DNA block is then digested with restriction enzymes usually *Sma*I or *Kpn*I for campylobacter, and loaded into the agarose gel and sealed in place. Once Electrophoresis starts the DNA migrates out of the block and into the agarose gel. Attempts have been made to standardise the protocols and results produced from PFGE so that fingerprints can be easily compared between laboratories. This has included the development of Campynet (Anon, 2001) in Europe and Pulsenet (Ribot *et al.*, 2001) in the US, which together with dedicated

software packages such as Bionumerics (Applied Maths) or Gel Compar (Applied Maths) has facilitated comparison of results between laboratories and use in identifying outbreaks.

1.18.2 Amplified Fragment Length Polymorphism (AFLP)

Amplified fragment length polymorphism is based upon the selective PCR amplification of restriction fragments from a total digest of genomic DNA. The technique involves three steps; (i) restriction of the DNA with one or two restriction enzymes and ligation of oligonucleotide adapters (ii) selective amplification of sets of restriction fragments, by using the adapter and restriction site sequence as target sites for primer annealing and, (iii) gel or sequence analysis of the amplified fragments (Vos *et al.,* 1995). The number of fragments that are visualised depends on the resolution of the detection system; typically 50-100 fragments can be detected by using denaturing polyacrylamide gels and 6-12 fragments by agarose gels. Amongst reported variations in the technique is an approach utilising a single enzyme digest and electrophoresis using agarose gels. This approach termed single enzyme AFLP (SAFLP) has been applied to *C. jejuni* and *coli* and other organisms including *H.pylori* (Gibson *et al.,* 1998), *C. perfringens* (McLauchlin *et al.,* 2000), *B.cereus* (Ripabelli *et al.,* 2000b), *S.typhi* (Nair *et al.,* 2000) and *L. monocytogenes* (Ripabelli *et al.,* 2000a).

Although all these methods harness the variation between different strains, they are reliant upon the variation in restriction sites, and not on the sequence variability between these sites. The majority of these methods (excluding FAFLP) involve the subjective interpretation of DNA migration on agarose gels. The use of computer

software such as Bionumerics to analyse banding patterns has made the process easier, however, time consuming normalisation needs to be applied in order for the standardisation of gels to allow comparisons between laboratories.

The techniques of PFGE and AFLP tend to be most widely used for subtyping campylobacter due to the high level of discrimination obtained. However the current methods for subtyping of campylobacter described are time consuming and tedious and suffer from limited standardisation, or recognised nomenclature making interlaboratory comparisons difficult. Additionally, very few of the techniques described can be applied to isolates from both a human source and a potential suspect source in a timely manner, making their use as tools in their present format for rapid epidemiological studies limited.

1.19 Sequence typing methods

Sequence typing methods are the next generation DNA typing methods and have only recently become a feasible option with the reduced cost and greater availability of sequencing platforms (Clarke *et al.,* 2001). Schemes have been described for campylobacter including the sequencing of the short variable region (SVR) of the *fla* A gene (Meinersmann *et al.,* 1997), Multiple Locus Variable Number Tandem Repeats Analysis (MLVA) (Hoffmaster *et al.,* 2002; Lindstedt *et al.,* 2003) partial 16S rRNA sequencing (Gorkiewicz *et al.,* 2003) and Multi Locus Sequence Typing (MLST) (Dingle *et al.,* 2001; Maiden *et al.,* 1998; Manning *et al.,* 2003). Sequence typing methods suit the study of bacterial population genetics and provide a snapshot of detail into the genome. They are especially useful for those bacteria such as campylobacter where comparatively little is known about transmission and

epidemiology due to absence of a universal and robust typing method with standardised nomenclature (Dingle *et al.*, 2001).

A substantial amount of the work within this thesis is based upon the sequence typing technique MLST for *C. jejuni* described by Dingle *et al* (Dingle *et al.*, 2001) which is described in the next sections.

1.19.1 The Campylobacter population structure

The campylobacter population structure has been described as weakly clonal (Dingle *et al.*, 2001; Dingle *et al.*, 2002; Manning *et al.*, 2003). In clonal population structures, variation accrues by intragenomic changes such as single nucleotide polymorphisms, rearrangements, deletions and insertions, which accumulate slowly over time (Kimura, 1991). Where populations are clonal, this may be attributed to the species being unable to accumulate sufficient sequence diversity from the ancestral strain due to the population being relatively young. For example, *Y. pestis,* has been described to have evolved from *Y. pseudotuberculosis* two thousand to fifteen thousand years ago (Achtman *et al.*, 1999). Some bacteria are clonal but have had sufficient time to accumulate sequence diversity, for example *S. enterica*, infects mammals and reptiles but is thought to have been derived from an ancestor which existed before mammals and reptiles evolved about one hundred and twenty million years ago (Achtman & Suerbaum, 2001). In weakly clonal populations such as campylobacter, the rate of intragenomic change is more frequent; therefore more variation is seen throughout the population. This is a highly desirable property for development of a typing scheme, where variation should make possible the characterisation of different campylobacter strains.

1.19.2 Multi Locus Sequence Typing (MLST)

MLST is based upon the well-tested principles of Multi Locus Enzyme Electrophoresis (MLEE) (Murphy, 1993) in that neutral genetic variation from housekeeping genes is indexed at multiple chromosomal locations. Alleles are assigned at each locus directly by nucleotide sequencing, rather than indirectly from the electrophoretic mobilities of their gene products (Enright & Spratt, 1999) as is the case for MLEE. Approximately 450-500bp internal regions of each of seven housekeeping genes are used, as these can be accurately sequenced in both forward and reverse directions using automated DNA sequencers (Dingle *et al.*, 2001). Currently, MLST strategies exist for various organisms, and established schemes have been set up for many different bacteria (Table 1.4), which can be accessed through the website http://pubmlst.org. This is one of the greatest advantages of MLST in that the centralized public databases are maintained and are extremely useful for the long-term concurrent data collection in independent laboratories.

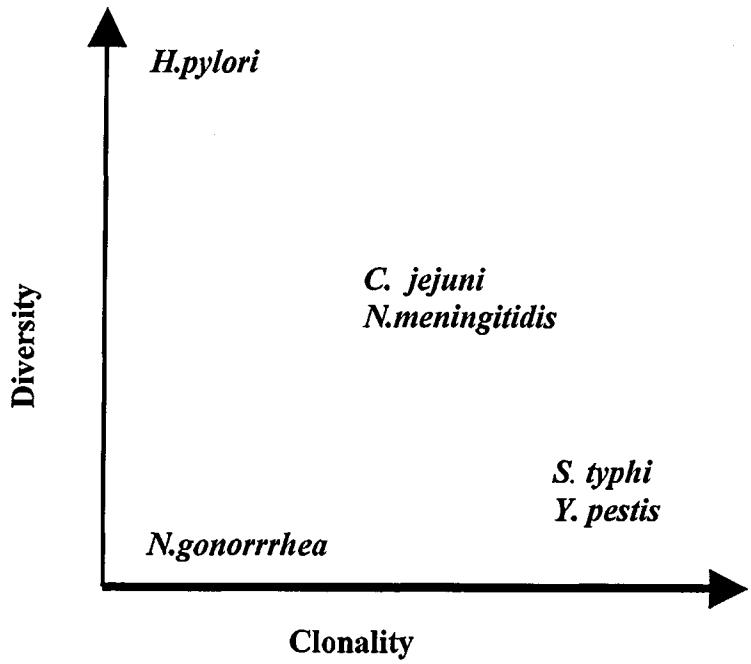
Table 1.4 Established MLST Schemes and references

Organism	**Reference**
Staphylococcus aureus	(Enright *et al.*, 2000)
Neisseria meningitidis	(Maiden *et al.*, 1998)
Streptococcus pneumoniae	(Enright *et al.*, 1998)
Campylobacter jejuni	(Dingle *et al.*, 2001)
Group B *Streptococcus*	(Jones *et al.*, 2003)
Enterococcus faecium	(Homan *et al.*, 2002)
Steptococcus pyogenes	(Enright *et al.*, 2001)
Vibrio cholerae	(Kotetishvili *et al.*, 2003)

Listeria monocytogenes	(Salcedo *et al.*, 2003)
Bartonella henselae	(Iredell *et al.*, 2003)

MLST strategies are not applicable for all bacteria and the utility of a sequence-typing scheme varies according to where the organism lies on a scale of variation versus recombination (Spratt, 2002). For example, a highly clonal organism with little divergence would have little sequence diversity, so would not be suitable for a MLST scheme, likewise a highly divergent species, but with low clonality such as *H.pylori*, would equally make a MLST scheme difficult to establish as too many different allele types which would be produced (Achtman & Suerbaum, 2001). Organisms with a suitable balance between clonality and divergence such as *C. jejuni,* which is described as having a weakly clonal population structure (Dingle *et al.*, 2002) appear to be the best type of organisms in which to apply MLST strategies (Figure 1.1).

Figure 1.1 Clonality versus diversity, the placement of organisms.



The analysis of nucleotide sequence variation in other organsims at multiple gene loci has permitted us to gain a further understanding of the population structure of a range of pathogens (Jolley *et al.,* 2000; Suerbaum *et al.,* 2001; Urwin & Maiden, 2003). These types of investigations have shown that different pathogens have widely differing sequence variation, population structure, rates of mutation and recombination, and also have demonstrated the existence of clonal grouping with respect to environmental locations (Colles *et al.,* 2003; Maiden *et al.,* 1998; Tzanakaki *et al.,* 2001). The suitability of *C. jejuni* for MLST and the success of the system for other organisms such as *N.meningitidis* have meant that MLST was anticipated to be a success in campylobacter typing, providing a potential strategy for universal typing, and overcoming the shortcomings of existing methods such as PFGE.

1.19.3 Bacterial House Keeping Genes

All bacteria possess a large number of genes, whose sequences are relatively conserved and whose products are vital to the normal functioning of a growing and dividing cell, regardless of the environmental conditions (Enright *et al.*, 2001). These genes are always active in growing bacteria and are known as constitutive or housekeeping genes (Russell, 2003). These genes include ones responsible for the basic functioning of the cell for example enzymes needed for protein and glucose metabolism. However this does not include the rRNA genes. Although rRNA genes are strongly conserved they are too conserved within species to be useful for a typing scheme. This is due to the fact that a change in the secondary structure of the rRNA that interacts with the ribosomal proteins is deleterious (Achtman & Suerbaum, 2001).

Housekeeping genes provide suitable targets for sequence typing schemes due to being present in every organism of the same species, and they accumulate mutations as a result of change over time (Russell, 2003). They are stable genes evolving at a slow rate under negative stabilising selection. When housekeeping genes from any organism are aligned they usually are continuous without any gaps, frameshift mutations or stop codons (Achtman & Suerbaum, 2001). Sequence variation in house keeping genes is also more likely to reflect the phylogeny of strains, as most of the mutations are non synonymous, unlike genes for example which encode immunogenic proteins where frameshift mutations or stop codons occur frequently, due to the selective pressure of the immune system (Achtman & Suerbaum, 2001). Virulence genes are affected by horizontal gene transfer and accumulate mutations rapidly under positive stabilising selection usually in response to environmental

stresses, resulting in unevenly balanced genetic diversity. Consequently, sequence analyses of virulence genes are much more likely to reflect information on host-parasite interactions rather than on the phylogeny of strains themselves (Spratt, 2002).

There is strong evidence that interspecies recombination occurs frequently enough in *C. jejuni* to create many different sequence types and this is clearly a major factor for generating genetic heterogeneity between strains (Dingle *et al.,* 2001; Suerbaum *et al.,* 2001). Analysis of recombination of this type has also been shown to occur in the *fla*A and *fla*B genes of *C. jejuni* where intragenomic and intrastrain variation has been shown to occur (Alm *et al.,* 1993; Harrington *et al.,* 1999; Park *et al.,* 2000; Wassenaar *et al.,* 1995).

1.19.4 Campylobacter *jejuni* MLST

Until the introduction of the MLST schemes there had been very few systematic analyses of nucleotide sequence variation in campylobacter. MLST is a method based on 400-500bp nucleotide sequences of seven housekeeping genes which has been shown to be a powerful technique for campylobacter typing (Colles *et al.,* 2003; Dingle *et al.,* 2001; Dingle *et al.,* 2002; Manning *et al.,* 2003; Matsuda *et al.,* 2003; Sails *et al.,* 2003b; Sails *et al.,* 2003c; Schouls *et al.,* 2003; Suerbaum *et al.,* 2001).

Three MLST schemes have been described for *Campylobacter jejuni.* The most widely referenced MLST scheme was published by Dingle *et al* (Dingle *et al.,* 2001) in January 2001, a similar MLST scheme was published by Suerbaum *et al* (2001) in April 2001 and a third scheme was described in November 2003 by Manning *et al*

(2003). These schemes use a selection of housekeeping genes from the published *C. jejuni* genomic sequence and were based on different house keeping genes. The Suerbaum scheme genes were selected based on the following criteria: they encoded housekeeping genes; they were widely separated on the chromosome and were not located in the vicinity of putative virulence genes or outer membrane protein genes. This strategy was tested on a selection of 32 *C. jejuni* strains from patients with enteritis from diverse origins including Germany, Hungary, Thailand and the US, also included in this strain set was the sequenced strain NCTC 11168. For all the loci analysed, multiple strains carried identical alleles and the frequency of synonymous and non-synonymous mutations was low. They found a total of 31 different combinations of allele types (sequence types), so that all but two strains could be distinguished from each other. The extent of recombination was analysed by the homoplasy test and split composition, which showed intraspecific recombination is frequent in *C. jejuni* and has generated extensive diversity of allelic profiles from a small number of polymorphic nucleotides. However one of the genes *atpA* used for this strategy was also present in *H. pylori*. Analysis of allele sequences showed that the *C. jejuni* and *H. pylori* sequences formed two distinct branches of the neighbour joining tree (Jukes-Cantor distances) with all the sequences within the 20 *H.pylori* strains being unique; there were only 9 allele types for the 32 *C. jejuni* strains analysed.

The genes used for Dingle MLST were chosen on the basis of similar criteria including chromosomal location, suitability for primer design, and sequence diversity in pilot studies. Seven loci (*asp*A, *gln*A, *glt*A *gly*A, *pgm*, *tkt*, *unc*A) (Table 1.5), were chosen for the MLST scheme for *Campylobacter jejuni* and as in the case of the

Suerbaum scheme the spacing of these genes around the genome (more than 70kb apart) means it would be highly unlikely for any two genes to be co-inherited in the same recombination event.

This system was established using 194 *C. jejuni* isolates from diverse origins including humans, animals and the environment. The data from this study confirmed that *C. jejuni* has a weakly clonal population structure, but was genetically diverse and that intraspecies and intrastrain recombination was common. Out of the 155 STs found, 51 (26% of the isolate collection) were unique and the remainder being categorised into 11 lineages or clonal complexes. The fact that this typing system seems to work for the amount of recombination in *C. jejuni* suggests that MLST is an applicable typing scheme for large-scale analysis of *C. jejuni* populations. The Suerbaum scheme has shown no relationships between MLST type and other methods such as serotyping, although the number of strains analysed would not permit associations of this type to be made. However, the Dingle scheme has shown certain associations with MLST clonal complexes and certain Penner serotypes (Dingle *et al.*, 2002).

The Manning *et al* (2003) scheme initially used six, then subsequently five loci instead of seven, with one (*glt*A) being the same as that used in the Dingle *et al* MLST scheme. The scheme was designed to investigate differences between 82 campylobacter isolates obtained from human and veterinary sources and to compare against the Dingle scheme. The authors confirmed the weakly clonal population structure of *C. jejuni* and showed that campylobacters from veterinary and human populations do overlap (Manning *et al.*, 2003). By comparison of the two schemes

they found that the largest clonal complex in each case shared the same isolates, but there were differences in the exact clustering, probably accounted for by the differences in the discrimination of the two schemes.

The Dingle Scheme has been widely utilised worldwide and is considered the main MLST scheme for *C. jejuni*. The success of the scheme can be mainly attributed to the website (http://pubmlst.org/campylobacter) which was developed for the scheme at Oxford University. This enables users to directly input sequences, to assign allele types, sequence types and clonal complexes ensuring standardisation. It also acts as a database for submission of data meaning that users can add to or compare their results with previously submitted data (Taylor & Fisher, 2003).

The Dingle MLST process (Dingle *et al.*, 2001) involves the amplification of a sequence of approximately 700-800bp of each of the seven loci (Table 1.5), with published primer sets followed by the sequencing of an internal 400-500bp region. An allele number is then assigned by interrogation of the *C. jejuni* MLST website. When an allele number for each of the seven loci, hence a seven digit code is obtained, the website is again utilised to assign the sequence type (e.g. ST-21), which is then assigned into a clonal complex (e.g. ST-21 clonal complex). When the scheme was set up all alleles were assigned arbitrary numbers in order of assignment, and new alleles are still assigned in order of recognition. Likewise all sequence types (STs) are assigned arbitrary numbers assigned in order of description e.g. ST-1 as the first. All STs were grouped into lineages or clonal complexes by use of the program BURST (E.J. Feil and M-S Chan http://mlst.zoo.ox.ac.uk). These were defined as groups of two or more independent isolates with an ST that shared identical alleles at

four or more loci, these were named after the founder member of the group e.g. ST-21 complex.

MLST provides a means for investigation of campylobacter epidemiology, which was considered difficult and complicated with the existing methods for typing. Whilst the MLST method is robust and reproducible it can be time consuming without robotic equipment and expensive initial set up, meaning its use could be limited in certain situations. However with the increasing use of laboratory automation (Clarke, 2002), the reduced cost of sequencing consumables (Diggle & Clarke, 2002) and confirmed inter and intra laboratory reproducibility of current sequencing methods (Clarke *et al.*, 2001; Diggle & Clarke, 2002), nucleotide sequencing boasts many attractive features which surpass other available methods for typing.

The wealth of information generated to date with MLST provides a resource for further study and scope for the further development of the technique. The MLST clonal complex has been established as an important epidemiological group at the strain level providing accurate and phylogenetically valid strain identification. Currently, twenty-four clonal complexes have been described, with ST-21 clonal complex being the largest, containing 26% of submitted isolates. Sixty-six percent of all submitted isolates assigning into one of six major complexes these being, ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257 (http://pubmlst.org/campylobacter). The data from preliminary MLST studies of *C. jejuni* isolates from animals and human cases of infection has established the concept of host associations between clonal complexes of *C. jejuni* such as poultry, and cattle, which also cause human

infections. For example, this has been described for clonal complexes ST-45 and ST-257 which are reported to contain isolates predominantly of poultry and human origins and ST-61 and ST-403 have been associated with bovine and ovine isolates but not human isolates (Dingle *et al.*, 2002; Manning *et al.*, 2003).

Conserved Single Nucleotide Polymorphisms (SNPs), within each of the MLST alleles, which identify the allelic profile of the major epidemiological lineages, have been identified from all the alleles within the current MLST database. Based upon this intelligence, it was considered that a parallel technique could be devised using the nomenclature and basics of MLST, but using an approach based upon the SNPs identified within MLST alleles, which were specific for the clonal complexes. This approach would require less time for a result, therefore be more beneficial for epidemiological studies.

Table 1.5 Detail of Housekeeping genes used in the Dingle *et al* (2001) *C. jejuni* MLST scheme.

Gene	Protein	Chromosomal Location	Role	Size (gene and protein)
aspA **Cj0087c**	Aspartase ammonia-lyase	96074-97480	Central Intermediary Metabolism	1407bp 468aa
glnA **Cj0699c**	Glutamine Synthetase	656901-658331	Amino Acid biosynthesis Glutamate family	1431bp 414aa
gltA **Cj1682c**	Citrate Synthetase	1603983-1605251	Energy Metabolism Tricarboxylic cycle	1269bp 422aa
glyA **Cj0402**	Serine hydroxymethytransferase	367219-368463	Amino Acid biosythesis Serine family	1245bp 414aa
pgm_ **Cj0434**	Phosphoglycerate Mutase	402285-403763	Energy metabolism glycolysis	1479bp 492aa
tkt_ **Cj1645**	Transketolase	1569190-1571088	Energy Metabolism Pentose Phosphate pathway-non oxidative branch	1899bp 632aa
uncA **Cj0105**	ATP synthase α subunit	111488-112993	ATP proton motive force	1506bp 501aa

1.20 The use of Single Nucleotide Polymorphisms (SNPs) for subtyping

In recent years a powerful approach involving the detection of single nucleotide polymorphisms (SNPs) has become increasingly popular in human genetics, the application of which has been extended to studies in bacteriology. By definition a SNP is a single point nucleotide change, in contrast to nucleotide insertions, deletions and small inversions. In the human genome SNPs are responsible for 98% of all polymorphisms (Russell, 2003) and have been associated with certain inherited disorders as well as susceptibility to infection and somatic diseases. By convention in human genetics the SNP must be present in 1% of the human population to be called a SNP, and occur every 100-300 bases, although no formal convention exists within bacteria (Russell, 2003). Nevertheless the importance of SNPs in bacterial genes that contribute to the ability of pathogens to cause disease or survive has been determined in many bacteria. SNPs leading to amino acid replacements can lead to niche expansion, and novel biological species (Perutz, 1983). There are two types of SNP:

- **Nonsynonymous polymorphism** is a single nucleotide polymorphism present in the coding region that alters the codon to result in an amino acid replacement.
- **Synonymous polymorphism** (Silent Mutation) is a single nucleotide polymorphism present in the coding region, which produces a synonymous codon, resulting in no difference to the polypeptide chain.

The advantages of using SNPs in both human and bacterial genetics can be attributed to a number of reasons. (i) They are a very abundant accessible class of polymorphism present in the genome of most species (ii) they are potentially easy to

detect and amenable to automation, therefore reducing costs and (iii) the inadequacy of other genetic methods, especially in bacteriology, to determine polymorphisms at a sequence level, other than with full sequencing or restriction enzyme analysis, makes them a more promising alternative.

Studies involving characterisation of SNPs in bacteria have been focused on those SNPs within structural genes, which can have a dramatic effect on the biology of whole organisms, and contribute to the ability of the organism to cause disease (Weissman *et al.,* 2003) or affect the susceptibility of an organism to particular agents (Ramaswamy *et al.,* 2003). Particular SNPs have been described in the *E. coli* fimbrial adhesin genes, which result in amino acid replacements associated with increased bacterial tropism for the uroepithelium and bladder colonisation (Sokurenko *et al.,* 1998). Likewise in *Salmonella typhimurium,* variation in SNPs within the type 1 fimbrial adhesin gene produce significant differences in cell binding, biofilm formation and host colonisation properties (Boddicker *et al.,* 2002). Informative SNPs have also been discovered within the genomes of the different strains of *Bacillus anthracis* (Read *et al.,* 2002). More recently SNPs have been identified to differentiate between lineages of *Listeria monocytogenes* (Moorhead *et al.,* 2003).

From the alleles defined by the Dingle MLST database, where 2162 profiles have been deposited, there are stable polymorphisms within the alleles, which identify the allelic profile of the major epidemiological lineages or clonal complexes and represent valuable markers for detection. With appropriate assay design, detection of the SNPs could define specific *C. jejuni* types (by clonal complex), which would not

only be recognisable by a defined nomenclature, but would also be directly comparable with the results already obtained through full MLST.

1.21 Background to Real Time PCR

The Polymerase Chain reaction, described by Panet and Khorana (1974), revolutionised the field of molecular biology, however the more recent advent of real time PCR techniques offers considerable further benefits. Real time PCR is the ability to monitor the progress of a PCR reaction as it occurs ("in real time"), where data is collected throughout the PCR process rather than at the end of the PCR. Many different real time PCR platforms are now available. Two different platforms for real time PCR are used within this study these being the Applied Biosystems Sequence Detection System (SDS 7000 or SDS 7700) more commonly referred to as the Taqman, and the Roche Lightcycler.

1.21.1 The Applied Biosystems Taqman

The principle for the Taqman chemistry was first described by Holland *et al* (1991). A method was devised whereby the simultaneous target amplification and generation of target specific signal were achieved by utilisation of the 5'-3' exonuclease activity of *Taq* polymerase. This activity of *Taq* DNA polymerase cleaves 5' terminal nucleotides of double stranded DNA releasing mono and oligonucleotides. The preferred substrate for optimal exonuclease activity is single stranded displaced DNA, in a fork like structure with the hydrolysis occurring at the phosphodiester bond joining the displaced region with the base paired portion of the strand. To facilitate this, a substrate (a probe) was created, suitable for exonuclease activity, which hybridised within the target sequence. During the PCR reaction the 5'-3'

exonuclease activity of *Taq* DNA polymerase degrades the probe in a nick-translation reaction and releases smaller fragments containing probe into solution (Applied Biosystems, 2001).

The Taqman system (SDS 7700) is a fully integrated system for real time PCR. The machine comprises a built in thermal cycler, a laser to induce fluorescence, and a CCD (charged couple device) detector connected to a computer which analyses and records the data. The newer instrument SDS 7000 is identical except for utilising a halogen light source instead of the laser to induce fluorescence. The Taqman system uses fluorogenic hydrolysis probes to enable the detection of a specific PCR product as it accumulates. The oligonucleotide probe is constructed with a reporter fluorescent dye on the 5' end (either FAM or VIC) and a quencher (either fluorescent or non fluorescent) dye on the 3' end (usually TAMRA). Whilst the probe is intact (outside of the PCR reaction) the proximity of the reporter dye to the quencher dye is sufficient so that the fluorescence emitted by the reporter dye is suppressed by fluorescence resonance energy transfer (FRET) through space. However during the PCR reaction, and if the target sequence is present the probe anneals to the sequence downstream from one of the primer sites and is cleaved by the 5' nuclease activity of *Taq* DNA polymerase as the primer is extended. At this point cleavage of the probe occurs and the reporter and quencher dyes are separated, hence increasing the signal of the reporter dye. The probe is removed from the target strand allowing the extension of the primers. The inclusion of the probe into the reaction does not adversely affect the PCR in any way. As the PCR reaction continues, subsequent PCR cycles result in further release of probe, which is proportional to the amount of amplicon present in the reaction. Thus Taqman assays are quantitative, and can be

used to measure the amount of nucleic acid target present. In the initial cycles of a PCR there is little change in the amount of fluorescent signal, which defines the baseline for the amplification plot; any increase in fluorescence above this baseline indicates that amplification has occurred and subsequent detection of the target sequence (Applied Biosystems, 2001).

This technique has completely transformed the PCR process, as reactions are characterised by the point in time during cycling when amplification of a target is initially detected, rather than the amount of target accumulated after a certain amount of cycles. For example the higher the starting amount of nucleic acid the sooner a significant increase in fluorescence is observed. Real time PCR offers rapid one-step PCR quantification within a closed tube, virtually eliminating any contamination, and allowing a huge reduction in hands-on time.

1.21.2 The Roche Lightcycler system

As in the case of the Taqman system the Lightcycler instrument consists of two main components, (i) a thermal cycler and (ii) a fluorimeter, which work together in the manner of the Taqman system. However the Lightcycler achieves rapid cooling and heating by a hot and cool air blower as well as the use of glass capillaries to contain the reaction mix. The high speed cycling in the Lightcycler is made possible by the design of these capillaries. They permit the reaction volume to be reduced to 10-20µl, and their optimised surface to volume ratio guarantees extremely rapid thermal transfer within the reaction mix. As a result, ramping times (the time taken to reach a required temperature) are very rapid and the total time taken for each cycle is approximately 15-20 seconds (Anon, 2003).

Two possible chemistry formats for detection are possible on the Lightcycler these being the (i) sequence independent fluorescent detection with the double stranded DNA binding dye SYBR Green I and (ii) sequence specific fluorescent detection with oligonucleotide probes that are coupled to suitable fluorophores.

(i) Sequence independent detection-SYBR Green I Dye chemistry

Small molecules which bind to double stranded DNA can be divided into two classes, intercalators and minor groove binders. However, in order to be successful within a real time PCR reaction they have to have the capability of increased fluorescence when bound to double-stranded DNA, and cause no inhibition to the PCR. The SYBR green dye detects polymerase chain reaction products by binding to double stranded DNA, which is newly formed during the PCR. The action of the *Taq* polymerase generates further double stranded amplicons, to which SYBR green binds. Subsequent PCR cycles result in an increase in amplicon number, therefore an increase in bound SYBR green and increase in fluorescence, which is proportional to the amount of DNA present (Anon, 2003).

(ii) Sequence dependent detection-Fluorescent hybridisation probes

For DNA detection and quantification studies the hybridisation probes format on the Lightcycler can be used which provides a maximal specificity for product identification. This is achieved though the use of two oligonucleotide primers as well as two further oligonucleotides (probes) which are labelled with fluorescent dyes and internal to the primers. The two probes are labelled in such a way that the first one (the anchor probe) is labelled with a fluorescein dye at its 3' end; the second probe

(the sensor probe) is labelled with a dye (either LCred 640 or 705). The two probes bind in close proximity during the PCR reaction; the fluorescein is excited by the Lightcycler's LED (light emitting diode) filtered light source, and the excited LC red 640 or 705 attached to the second hybridisation probe, subsequently emits light at a longer wavelength than the fluorescein. The FRET occurring is highly dependent upon the spacing of the two probes, or their dye molecules, only if they are in close proximity (approximately 1-5 nucleotides apart) can energy be transferred at high efficiency. With the correct choice of detector channel (depending on which dye combination is used) the intensity of the light emitted by the Lightcycler red 640 or 705 is filtered and measured by the Lightcycler Instrument's optics (Anon, 2003).

The Lightcycler system allows for the simultaneous detection of two different target sequences in one sample, which has numerous applications for genotyping and detection purposes. This allows the user to obtain more results from a single reaction; more mutations or specific sequences can be obtained within one run. The Lightcycler instrument is equipped with a three-channel fluorimeter, which detects emitted light with wavelength maxima of 530, 640, and 705 nm. Channel 1, which detects at 530nm, can be used for the detection of dyes such as SYBR Green I and Fluorescein. Channel 2, which detects light emitted at wavelengths of 640nm, is optimised to detect the Lightcycler dye Lightcycler Red 640 (LCred 640), and likewise channel 3 which detects light emitted at wavelengths of 705nm is optimised to detect the light emitted by the dye Lightcycler red 705 (LCred705) (Anon, 2003).

1.23 Aims of the Study

Campylobacter jejuni and *Campylobacter coli* are the major causes of bacterial enteritis in the developed world, yet transmission routes and epidemiology are far from clear. Although many typing schemes have been devised for campylobacter very few are applicable for the rapid identification of isolates to strain level and are applicable both to isolates from culture and for the detection and typing of isolates within other matrices such as food and environmental sources. The rapid identification of *C. jejuni* isolates to a recognised strain level directly from potentially contaminated samples would both inform and enhance the epidemiological investigation of *C. jejuni.*

The aim of this study was to devise strategies for rapid detection and identification of *C. jejuni* from potential sources of campylobacter, which could be applied to food and environmental samples. Since many of the described methods for *C. jejuni* typing are not applicable for the rapid identification of recognised strains from matrices other than culture, the majority of the study concerned the development of improved methods for specific strain identification as soon as possible in the investigation process.

A crucial factor in the development of a strategy for the detection of campylobacter is the ability to detect specific strains of campylobacter. A requirement for this would be the ability to recognise campylobacter subtypes by universally recognised nomenclature. As described, the *C. jejuni* MLST scheme provides a discriminatory molecular profile, is reproducible with simplicity of interpretation and provides data, which is directly comparable between laboratories via the Internet. For these reasons

the strategy proposed for this project was based upon the detection of specific campylobacter subtypes utilising the SNPs identified within the alleles of the Multi Locus Sequence Typing scheme clonal complexes.

This investigation involves:

- The development and testing of the applicability of using whole genome molecular fingerprinting techniques for identification of campylobacter.

- The development and evaluation of a real time PCR speciation assay for the identification of *C. jejuni* and *C. coli* species and application to artificially contaminated and naturally contaminated food samples.

- Extensive analysis of the existing MLST data, then development and evaluation of rapid real time allelic discrimination assays to identify strain associated single nucleotide polymorphisms (SNPs) based on MLST Locus alleles, which potentially inform clonal complexes. Application of the above technology to naturally contaminated food and water samples.

Chapter 2

Materials and Methods

Chapter 2 Materials and Methods

This chapter describes the general methods used throughout the study, more specific methods are described in detail in the relevant chapters.

2.1. Bacterial strain collections

Strains were primarily those received into the Campylobacter Reference Unit (CRU), and were cultured following frozen storage. Specific sets included the Multi Locus Sequence Typing (MLST) reference collection (Wareing *et al.*, 2003), received from Professor A.J. Fox, Health Protection Agency North West, Manchester, a collection from the Campynet project (www.campynet.dk) and a strain collection received from Mr. K. Stapleton, Veterinary Laboratories Agency, Weybridge.

2.2 Non culture samples –Food and water

All chicken, turkey, beef, lamb, pork and duck samples were purchased from various retail establishments in North London within a six-month period. Water samples were obtained from Dr S. Surmann-Lee and Dr S. Lai, Food Water and Environmental Laboratory, Health Protection Agency, London.

2.3 Culturing Techniques

2.3.1 Isolates from frozen storage

Materials

- Cryobeads and vials (Microbank; Prolab Diagnostics, Ontario, Canada)

- Columbia Blood Agar (CBA) (Oxoid CM0689, Basingstoke, UK.) with 5% defibrinated horse blood.
- Anaerobic gas jars (Don Whitley Scientific, Shipley, UK.)

Method

Cultures were stored on cryobeads at –80°C. One bead was removed onto a CBA culture plate. The bead was used to make a pool of bacteria on the surface of the agar, and then discarded. The plate was inoculated using standard techniques with a sterile 10µl culture loop to isolate single colonies. Inoculated plates were incubated at 37°C for 24 hours in a Variable Atmosphere Incubator (VAIN) under microaerobic conditions (5% CO_2, 5% O_2, 3% H_2, and 87% N_2).

2.3.2 Isolates received on charcoal transport swabs

Campylobacter isolates were referred to the CRU in Aimes Transport Medium with charcoal (Oxoid TS0002) as swabs. These are designed for the transport and maintenance of cultured cells, allowing for easier transportation by post.

Method

The transport medium was either subcultured as described above or subjected to direct DNA extraction. Direct extraction from the transport swab medium required the preparation of homogenous 1.2 diluted suspensions of the swab medium in nuclease free water. These were vortexed and the DNA extracted as described below (section 2.4.4).

2.3.3 Naturally contaminated meat samples

Method

All meat samples were tested within 24 hours of collection and stored as necessary at 4°C. The method used was an adaptation of the technique described by Kramer *et al.* (2000). Samples (5g) were weighed into 15ml campylobacter enrichment broth within a 20ml universal with the top secured, and shaken vigorously for two minutes to release the cells into solution. After leaving the contents to settle 100µl of the rinse fluid was either used directly in DNA extractions or removed and plated onto CCDA. CCDA plates were incubated under microaerobic conditions at 42°C for 48 hours. The remainder of the rinse fluid was incubated at 42°C in microaerobic conditions for 48 hours, plated (100µl) onto CCDA and incubated as before.

2.3.4 Enumeration of bacteria for spiking experiments

Materials

- Nuclease free water (Promega, Southampton, UK)
- CBA plates (Oxoid)

Method

This method was adapted from Standard Operating Procedure F-7338/03-99 written by K Pathak, Food and Water Environmental Laboratory, Colindale, London. Triplicate 24-hour identical cultures of *C. jejuni* were resuspended in nuclease free water to an optical density of 0.3 OD_{600}. Ten-fold serial dilutions from 1 to 10^{-9} were prepared and the inoculum levels determined by surface plating 5µl of each dilution onto CBA plates. These were incubated as normal, the colonies counted and mean numbers calculated.

2.3.5 Preparation of meat rinse samples and artificial contamination

Materials

- 5ml sterile bijoux (Alpha Laboratories, Eastleigh, UK)
- Campylobacter Enrichment Broth (Bolton Broth, HPA Media Dept, Colindale, London. UK)
- Campylobacter Selective Blood Free agar CCDA (Oxoid PO0119)

Method

Samples (5g) were shaken vigorously in a 20ml sterile universal with 15mls campylobacter enrichment broth for 2 minutes. Viable campylobacter cells were determined by inoculation onto CCDA and parallel enrichment in campylobacter enrichment broth and subculture onto CCDA. Suspect colonies from the CCDA pates were purified by subculture onto CBA. For spiking experiments a meat rinse suspension from a sample with no detectable campylobacter by standard culture (CCDA and CBA) was removed and divided into ten 900µl aliquots. 100µl of each of the standard suspensions of *C. jejuni* cells was added producing a range of dilutions from 10^{-1} to 10^{-10}.

2.4 DNA extraction methods from cultured cells

Various methods for DNA extraction were used throughout this study. This ranged from the crude DNA preparation by the use of cell lysates to the use of the more complex CTAB extraction method. Newer methodologies became available throughout the course of the project including the Isoquick Kit and the MagNApure instrument for automated DNA extraction.

All DNA extractions were carried out using culture plates incubated for 24 hours. The cells were removed with a moistened cotton swab and resuspended in 500µl nuclease free water.

2.4.1 Preparation of cell lysates

Crude DNA preparations were prepared using a single touch of a colony emulsified in 100µl distilled water. This was heated for 10 minutes at 100°C in a heating block and used directly in PCR reactions.

2.4.2 DNA Isolation using the hexadecyltrimethyl ammonium bromide (CTAB) method.

Most bacterial DNA extraction procedures consist of lysozyme or detergent lysis followed by incubation with a non-specific protease and phenol/chloroform extractions before alcohol precipitation of the nucleic acids. These procedures are effective in removing proteins but not effective in removing the contaminating polysaccharides. The CTAB solution combines with any residual exopolysaccharides and proteins that can be later removed by the phenol/chloroform extraction. DNA is isolated which is free from the copious amounts of exopolysaccharide produced by many bacteria, which can interfere with molecular biology reagents (Wilson *et al.*, 1987b).

Materials

- TE buffer (pH 8): 12.11g 10mM Tris (Sigma), 3.72g 1mM EDTA (Sigma), dissolved in 900µl distilled water, made up to 1 litre and diluted 1 in 10.

- CTAB Solution: 4.1g NaCl in 80ml water, slowly added 10g CTAB (hexadecyl trimethyl ammonium bromide). Heated to 65°C to dissolve. Adjusted volume to 100ml (stored at 25°C) (Sigma)
- 24:1 chlorofrom/isoamyl alcohol: 24ml chloroform, 1ml pentan-2-ol (Prepared in a fume hood) (Sigma)
- Microcentrifuge (IEC Micromax)
- phenol/chloroform/isoamyl alcohol (Fluka Biochemika)
- Freeze drier (Edwards Pirani 11)
- NaCl 5mM (Sigma)

Method

A saturated suspension (MacFarlane 5) of campylobacter cells was prepared in 900µl 0.85% saline and centrifuged in a covered rotor microcentrifuge for 5 minutes at 13000rpm. The resulting pellet was resuspended in 500µl TE buffer and 5µl Proteinase K (25mg/ml), 30µl 10% Sodium Docedyl Sulphate was added, and the suspension incubated at 37°C for 1 hour. 100µl of 5M NaCl was added and mixed thoroughly to avoid the formation of a CTAB nucleic acid precipitate (if the salt concentration fell below 0.5M). 80µl of CTAB solution was added and incubated at 65°C for 10 minutes. The aim was to remove protein and polysaccharides complexed to CTAB, leaving the DNA in solution.

An equal volume of chloroform/isoamyl (24:1) alcohol was added and mixed thoroughly to give even turbidity, and then centrifuged for 5 minutes (13000 rpm). The aqueous, viscous upper layer containing DNA was transferred into a fresh tube,

leaving the interface behind. An equal volume of phenol/chloroform/isoamyl was added and the tubes centrifuged as before. The supernatants were transferred to fresh tubes and 300µl isopropanol was added. The tubes were placed onto multimix rollers to precipitate the DNA, which was collected by centrifugation (13000rpm, 2 minutes). The isopropanol was removed and the pellet washed with 70% ethanol to remove the residual CTAB. The DNA was freeze-dried for 5 minutes and then redissolved in 100µl distilled water. The optical density of the DNA in the sample was calculated using a Biophotometer (Eppendorf, Fisher Scientific, Loughborough.UK).

2.4.3 DNA isolation using the Isoquick™ Kit

The Isoquick Kit utilises the chaotropic properties of guanidine thiocyanate (GuSCN), which both disrupts cellular integrity and inhibits nuclease (DNase and RNase) activities, in doing so provides lysis and DNA stabilisation. The GuSCN lysate is then mixed with a non-corrosive extraction reagent by centrifugation with the nucleic acid portioning into the aqueous phase. The nucleic acid is precipitated with alcohol, and then dissolved in RNase free water. The procedure was carried out at room temperature allowing multiple samples to be extracted simultaneously and without the use of corrosive volatile or highly toxic reagents.

Materials

Isoquick Kit (Orca Research), Microcentrifuge (IEC Micromax)

Methods

Extractions were carried out according to manufacturer's instructions for rapid DNA Extraction. Cells in an enrichment broth or other cell suspension were harvested by centrifugation and resuspended in reagent A (Sample buffer) to between 10^8-10^{10} gram negative bacteria per ml and incubated for 10 minutes at room temperature. The cellular sample (100µl) was lysed and stabilised by the addition of 100µl reagent 1 (Lysis solution) into a microcentrifuge tube. Reagent 2 (Extraction Matrix) was shaken and 700µl was added into each sample followed by 400µl of Reagent 3 (Extraction Buffer) and each sample vortexed for ten seconds and centrifuged (12,000rpm) for five minutes. The aqueous phase sample was transferred into a new microcentrifuge tube the volume estimated and 0.1 volume of reagent 4 (sodium acetate) was added to the aqueous phase sample. An equal volume of isopropanol was added to the contents of each microcentrifuge tube, and the tubes mixed gently to precipitate the nucleic acid. The samples were centrifuged (12,000rpm) for ten minutes and the supernatant discarded, without discarding the nucleic acid pellet. One ml of 70% ethanol was added the tube and then centrifuged (12000rpm) for 5 minutes. The supernatant was again discarded and the pellet allowed to dry. The DNA was re-suspended in 100µl RNAse free water and the concentration calculated as before.

2.4.4 DNA isolation using the Roche™MagNApure

The Roche MagNApure LC instrument is a robotic workstation for fully automated nucleic acid isolation from a range of different sample types (Figure 2.1). It can also be used for the setting up of Lightcycler capillaries for PCR reactions. Up to 32 samples can be processed for DNA or RNA (or both) from sample types such as blood, plasma, serum, cells, tissue, swabs or stool samples.

The system extracts the nucleic acid from 32 samples within 60-90 minutes (depending on the number of samples) within a closed housing, making it a completely walk away system. Different kits can be used on the machine depending on the type of extraction matrix and preferred isolation (e.g. total Nucleic Acid, DNA or RNA). All the extraction kits are based on the DNA isolation methods of Boom *et al.* (1990).

Two different kits were used on the Roche MagNAPure instrument, these being the Total Nucleic Acid extraction Kit III and the Bacterial DNA Isolation Kit. For both kits the isolation principle was based upon magnetic bead technology, the samples were lysed by incubation with the supplied buffer containing a chaotropic salt and proteinase K. Magnetic glass particles were added binding the nucleic acids, and remaining unbound substances removed by washing steps (Figure 2.2).

Figure 2.1 The MagNApure Instrument (Roche Applied Science 2003)

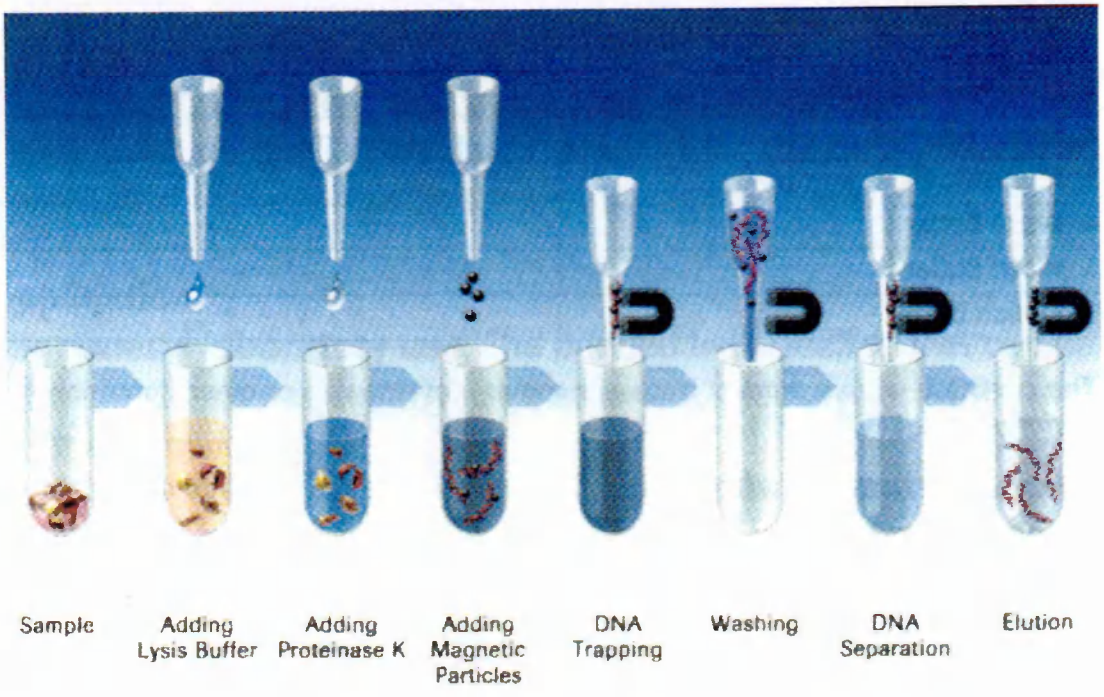


Figure 2.2 The MagNApure extraction process (Roche Applied Science 2003).

MagNAPure LC Total Nucleic Acid Isolation Kit I

Materials

- Roche MagNApure Total Nucleic Acid Isolation Kit III.
- Proteinase K, made up with 5ml of Elution buffer (provided in the kit)

Method

Using a 24-hour plate of culture, a saturated suspension was prepared in 500µl nuclease free water in a 2 ml microcentrifuge tube. This was vortexed to remove lumps and 200µl per sample was added into the MagNAPure sample cartridge. Before use the MagNAPure instrument was decontaminated using DNA Zap. The machine was set up with all the required plasticware, tips and reagents as specified by the manufacturers instructions and set to run on the required protocol. On completion, all the plasticwares and reagents were discarded and the machine decontaminated.

MagNApure LC DNA Bacterial DNA Isolation Kit III (Bacteria, Fungi)

This kit was specifically designed to isolate high quality intact DNA from samples such as stool, urine and sputum and therefore was used in this study for extraction from enrichment broths and meat rinses. The samples were first lysed with the Bacteria Lysis Buffer and Proteinase K, then inactivated at 95°C for 10 minutes and 65°C for ten minutes prior to transfer into the MagNApure Sample Cartridge. The buffer containing a chaotropic salt and magnetic glass particles to bind the DNA was added. All unbound substances were removed by washing steps as before.

Materials

- MagNApure LC DNA Bacterial DNA Isolation Kit
- Proteinase K made up by addition of 2.7ml of elution buffer

Method

A 24-hour culture plate was used to make up a saturated suspension as before or 100µl of enrichment broth of meat rinse sample was used. This was added to 130µl of bacterial lysis buffer and 20µl Proteinase K, incubated at 65°C in a water bath for ten minutes followed by a further incubation using a heating block at 100°C for ten minutes. The complete lysed sample solution was added to the sample cartridge and loaded into the MagNApure machine. The machine was set up as described previously but following the manufacturers guidelines for the Bacterial DNA Isolation Kit.

2.4.5 Archiving of DNA on FTA filters

A commonly used method in forensic laboratories is the storage of genomic DNA on FTA cards. These are described as an efficient system for the long-term storage of DNA, which have been described to be effective within most processes where purified DNA would usually be utilised including PCR, RFLP and RT-PCR. The FTA card is composed of a sophisticated filtration matrix impregnated with a patented formulation of powerful protein denaturants, a chelating agent and a free radical trap designed to protect and entrap nucleic acids (www.whatman.com). The formulation is non-toxic to humans, but prevents the overgrowth of other organisms

and allows the nucleic acid to be stored on the cards for many years without any reported degradation (Hansen & Blakesley, 1998).

Materials

- FTA Purification Reagent (Whatman, Fisher Scientific, Loughborough.UK)
- TE buffer (Whatman)
- Micropunch tool and mat (Whatman)

Method

Loading of FTA cards

A suspension of campylobacter (MF5) cells was prepared and 200µl was added onto the centre of the FTA card and allowed to air dry for 10 minutes. When dry, the cards were stored at room temperature for six months.

FTA card preparation for PCR

A 1.2mm diameter punch was taken from the FTA card and placed into a micro centrifuge tube containing 200µl FTA purification reagent, this was vortexed and incubated for 5 minutes at room temperature. The wash step was repeated twice then TE was used for a further two washes. All traces of liquid were removed with a pipette and the punch was allowed to air dry at room temperature for 1 hour. When dry, the punch was ready to add directly into a PCR reaction.

2.4.6 RNAse treatment of DNA Samples

Materials

- Ribonuclease A (Roche) (10mg added into 10 ml of distilled water and boiled for ten minutes).

Method

For accurate construction of standard curves and sensitivity studies isolated DNA was RNAse treated to prevent the over estimation of the quantity of DNA due contaminating RNA. DNA was treated with 1μg RNase preparation and heated in a thermocycler at 37°C for 30 minutes.

2.5 Non Culture DNA Isolation Methods

2.5.1 "Bugs 'n Beads" DNA Isolation

The "Bugs 'n Beads" Kit is a magnetic bead system for the isolation of PCR-ready DNA, it is not designed for the isolation of bacteria. The system uses magnetic beads, which have a bacteria-binding surface; these are mixed with the food sample containing the bacteria. Following a short incubation of 5 minutes, which allows time for the bacteria to bind to the beads, the DNA is extracted from the debris left in the solution by applying a magnet to the side of the tube. The remaining DNA/bead complex is washed several times to remove any cell debris, leaving PCR ready DNA.

Materials

- 96% and 70% Molecular grade ethanol (Sigma)
- "Bugs 'n Beads" Kit (Web Scientific, Crewe, UK.)
- Magnetic Separator (Web Scientific, Crewe UK)

Method

20µl (7.5mg/ml) of bacteria binding beads were added to 800µl BW buffer into a 1.5ml centrifuge tube, 100µl of the sample was added and incubated at room temperature for 5 minutes. The tubes were placed into the magnetic separator and after allowing time for the bacteria/bead complex to move towards the side of the tube, the supernatant was removed using a sterile pipette. The magnet was removed; the beads were resuspended in 50µl lysis buffer, and mixed by shaking the tube briefly. The samples were incubated at 80°C for 5 minutes with the lids closed using a heating block. 150µl of refrigerated 96% ethanol was added to the tubes and the incubation continued at room temperature for a further 5 minutes. The magnet was then re-combined with the tube and after allowing for the beads to move to the side of the tube, again the supernatant was removed and discarded. The magnet was separated from the tubes again and the beads washed with 70% ethanol. The magnet was replaced and the supernatant removed and discarded. This washing step was repeated once more, and all residual ethanol was removed. The DNA/beads complex was resuspended in 30µl nuclease free water and incubated at 80°C for 5 minutes with the lids open to remove the final residual ethanol.

2.5.2 Prepman Ultra® DNA Isolation (Applied Biosystems)

The Prepman Ultra reagent is a commercial product designed for the fast and convenient preparation of DNA from gram-negative pathogens present in food samples. It is a novel and proprietary clear homogenous solution and does not contain chelex or any other type of resin, matrix or suspended material. It potentially offers a less tedious and faster method for DNA isolation with lower costs. Little

information is available about the nature of the Prepman reagent activity, however it is based on template preparation by the inactivation of inhibitors such as nonionic detergents that denature proteins present in the sample. Clearly the resulting DNA preparation is not as pure as that obtained from a standard DNA extraction procedure from culture, but the literature describes the Prepman treated product to be of sufficient quality for most PCR applications (Applied Biosystems Prepman Kit Insert).

Materials

Prepman Ultra Sample Reagent (Applied Biosystems), Microcentrifuge (IEC micromax)

Method

The sample was aliquoted into a 1.5ml microcentrifuge tube and centrifuged at 13000rpm for 2 minutes to pellet the bacteria. The supernatant was removed and the remaining pellet used for extraction.

Before the Prepman reagent was used, the bottle was carefully swirled to ensure the liquid appeared uniform without any phase variation, and the volume required for use was aliquoted into a separate container to prevent contamination of the remaining liquid. 200µl of the Prepman reagent was aliquoted into the tubes containing the pellet, and the tubes capped and the pellet resuspended by vigorous mixing. The tubes were then incubated at 100°C for 10 minutes using a heating block and left to cool for 2 minutes at room temperature before centrifuging at 13000xg for 2 minutes. Following the centrifugation the supernatant was removed and diluted 1:2.

2.6 Methods for Serotyping and Phage typing

All serotyping and phage typing described in this thesis was carried out by the staff within the Campylobacter Reference Unit by published methods (Frost *et al.,* 1998; Frost *et al.,* 1999).

2.7 Molecular Fingerprinting Methods

2.7.1 Pulsed Field Gel Electrophoresis (PFGE)

This was carried out according to the Laboratory of Enteric Pathogens (Health Protection Agency. UK.) Standard Operating Procedure E-3221 for PFGE with the restriction enzyme *Sma*I.

Materials

- TBE buffer: 44.5 mM Tris-Borate and 1 mM EDTA. (Sigma)
- TE buffer: 10 mM Tris and 10 mM EDTA, pH 7.5. (Sigma)
- Proteinase K (50 mg/ml; Sigma)
- Lysis buffer: 1% w/v N-laurylsarcosine (BDH Laboratory Supplies) and 0.5M EDTA (Sigma), pH 9.5
- Saline (0.85% w/v NaCl)
- Low Gelling Agarose (Sigma type VII)
- Agarose (Sigma Type II: Medium EEO)
- *SmaI* and reaction buffer 4 (Invitrogen)
- CHEF DR-II (Bio Rad labs, USA)

Method

A bacterial suspension equivalent to MF standard 5 was prepared in 2.0ml saline; 0.45ml was mixed with 0.45 ml of 2% w/v low gelling agarose and dispensed into block moulds (10 mm x 5 mm x 1 mm). The gel was allowed to set within the blocks for 30 minutes at 4°C. The blocks were transferred into bijoux containing 2ml lysis buffer and 20µl Proteinase K and incubated overnight in a 56°C water bath. The lysis mix was then replaced with 2 ml TE buffer The TE buffer was decanted and replaced with fresh, TE buffer and incubated at 56°C for 30 minutes. This rinsing step was repeated twice then the liquid replaced with 2ml of fresh, cold TE buffer prior to cutting.

Gel blocks were removed from TE buffer and a portion was cut to 3.5 mm x 5 mm. The remaining block was replaced in the TE buffer and stored. The cut blocks to be digested were placed in 100µl of reaction buffer 4 for 30 minutes in sterile micro-centrifuge tubes at room temperature. The reaction buffer was decanted and replaced with 100µl of fresh buffer containing 20 units of *Sma*I and digested for 6 hours at 25°C. Digested blocks were loaded into preformed wells of a 1 % agarose gel and molecular weight markers (λ ladder, Bio Rad) were added to the outside and centre lanes. The loaded wells were sealed with molten agarose and the gel was loaded into the PFGE tank filled with 1750ml 0.5 x TBE buffer cooled to 14°C. *C. jejuni Sma*I DNA fragments were separated by PFGE at 6 Volts/cm for 22 hours with an initial pulse time of 10 seconds and a final pulse time of 35 seconds.

The DNA fragments were visualised by staining the gel with 1mg/ml ethidium bromide (for 1 hour, followed by destaining in distilled water for 2 hours) and

transilluminating with UV light (UVP Dual Intensity Transluminator). Restriction fragment migration profiles were compared and contrasted using the computer program Bionumerics (Applied Maths, Kortrijk, Belgium.). A single band difference was used to categorise a new profile type.

2.7.2. Single Enzyme Amplified Fragment Length Polymorphism (SAFLP)

Materials

- Digestion: *Hin*dIII and reaction buffer (Invitrogen), Spermidine (Sigma)
- Ligation: T4 DNA Ligase and 5x ligase buffer (Invitrogen), Adapter ADH1 (0.1) (5' ACG GTA TGC CAC AG 3'), Adapter ADH2 (0.1) (5' AGC TCT GTC GCA TAC CGT GAG 3') (MWG Biotech).
- Amplification: *Taq* DNA polymerase, 10x reaction buffer, 2.5mM $MgCl_2$ (Invitrogen), Primer HI-C (5' GGT ATG CGA CAG AGC TTC 3') (MWG Biotech).
- PTC200 Peltier Thermal Cycler, (MJ Research, Braintree, UK)

Method

DNA was obtained from 24-hour cultures using the CTAB (hexadecyl trimethyl ammonium bromide) or Isoquick technique. The concentration of DNA was estimated using a Biophotometer and made up to a concentration of 0.266mg/ml. The SAFLP method was carried out as described by Gibson *et al* (Gibson *et al.,* 1998).

DNA was digested with the restriction endonuclease *Hin*dIII in a total volume of 20µl containing 1µl of 0.1M spermidine 2µl 10x buffer and 2µl *Hin*dIII at 37°C

overnight. Ligation was carried out for 3 hours at room temperature in a total reaction volume of 20µl containing 5µl digested DNA, 2µl distilled water, 4µl 5x ligase buffer, 4µl each adapter and 1µl T_4 DNA ligase. The digested ligated DNA was heated to 65°C for 10 minutes to inactivate the ligase and subsequent PCR was carried out in a amplification mix of 50µl containing, 200µM of each dNTP, 2.5mM $MgCl_2$, 300ng primer, 0.2 µl *Taq* polymerase in 10X PCR buffer with 5 µl digested, ligated DNA. PCR conditions used were 94°C for 4 minutes for 1 cycle followed by 94°C for 1 minute, 60°C for 1 minute and 72°C for 2.5 minutes for 33 cycles. PCR products were detected on a 1.5% agarose gel run at 100v for 3 hours, followed by staining in 1mg/ml ethidium bromide for 20 minutes. SAFLP profiles were visualised under UV light, photographed and analysed both by eye and using the computer program Bionumerics (Applied Maths, Belguim).

2.7.3 SAFLP Method (adaptations to original method)

The digestion and ligation stages of SAFLP were combined in one reaction mixture. A mastermix was made up containing per 25µl reaction 2U *HindIII,* 2µl 10x React buffer,1U T4 DNA ligase, 4µl 5x ligase buffer, 0.2µg adapters ADH1 and ADH2 and 8µl nuclease free water. This was incubated at 37°C for 2 hours. Following this the samples were heated for 65°C for 10 minutes to inactivate the ligase and PCR was carried out as previously described.

2.8 Bionumerics Software

The Bionumerics software (Applied Maths) was designed to manage most biological experimental data, although it is advanced in the analysis of gel fingerprint patterns produced from subtyping techniques such as PFGE and SAFLP (www.applied-maths.com). It enables standardisation of results, as all gels can be analysed in the same way and provides a database where results can be stored alongside experimental data.

All the gel images produced either by PFGE or SAFLP were saved as TIF files and opened in the Bionumerics software. Using the software all fingerprint gels were analysed by standard methods and all normalised using the same parameters. This enabled dendrograms to be created to show the relatedness (or non relatedness) between different strains by the degree of homology between the fingerprints.

2.9 Multi Locus Sequence Typing (MLST) of *C. jejuni*

Two variations in the MLST method had to be used due to availability of sequencing instruments, the ABI Prism 3700 and the Beckman CEQ8000. Both methods used the same initial PCR amplification step however the subsequent steps differed due to the different sequencing chemistries used on the two machines. There are published protocols for MLST on the ABI instrument (Dingle *et al.*, 2001) (www.mlst.net/campylobacter). The Beckman CEQ8000 uses different chemistries, and an adapted method was therefore used.

2.9.1 Amplification PCR (Both Sequencing Platforms)

Materials

- *Taq* DNA Polymerase, Q solution, 10X buffer (with 15mM $MgCl_2$) (Qiagen)
- Primers-7 forward and 7 reverse (Table 2.1) (MWG Biotech)
- Deoxynucleotide triphosphate Mix (dNTPs) 10 mM (Applied Biosystems, Warrington, UK)

Method

Seven amplification PCR reactions were carried out per isolate to be typed by MLST. Each reaction was carried out in a total volume of 50µl containing per reaction 34.5µl distilled water, 5µl 10x PCR buffer, 5µl 5x Q solution, 1µl 15mM dNTP's, 1µl each of forward and reverse primers and 0.5µl *Taq* polymerase. 48µl of the above mastermix was added to 2µl genomic DNA (neat MagNApure extract). Cycling conditions were as follows 95°C for 3 minutes, then 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, and 72°C for 1 minute. Followed by a final soak of 10 minutes at 72°C. After cycling, products were run on a 1% agarose gel to check for correct size of product.

2.10 Sequencing Method 1- based on the Beckman CEQ 8000 Capillary Sequencer

2.10.1 PCR Clean -Up (Using Millipore Multiscreen Plates)

The Millipore Multiscreen plates using vacuum filtration provided a fast method for high throughput purification of PCR products. The PCR product is placed onto the filter in the well. On application of a vacuum all small molecular weight products i.e.

unincorporated dNTPs, primer and other PCR reagents, are pulled through the filter, leaving the large PCR products, which are too large to pass through the filter behind. By adding nuclease free water and agitating the plate the PCR products are resuspended and sufficiently pure for use in sequencing reactions (www.millipore.com/publications.nsf/docs).

Materials

- Nucleotide free water (Promega)
- Multiscreen PCR Plate and adhesive lid (Millipore, Watford, UK)

Methods

50µl of PCR product from above was loaded into each well of the multiscreen plate, the plate was placed onto the vacuum manifold and a vacuum applied at 17inches Hg for 10 minutes, or until the wells emptied. 30µl of nuclease free water was added into each well and the plate put onto an orbital shaker for 10 minutes, to resuspend the DNA. The DNA was then aliquoted into 0.5ml microcentrifuge tubes and stored at 4°C.

2.10.2 Sequencing PCR reaction

Materials

- Dye Terminator Cycle Sequencing (DTCS) Quick Start Kit (Beckman Coulter)
- Sequencing Primers -7 forward and 7 reverse (Table 2.2) (MWG- Biotech)
- Thermofast PCR 96 well plates and adhesive lids (AB gene)
- Nuclease free water (Promega)

- PTC200 Peltier Thermal Cycler, (MJ Research, Braintree, UK)

Method

Sequencing PCR reactions were carried out in Thermofast 96 well plates. Each 7.5µl reaction contained 2µl distilled water, 2.5µl DTCS mastermix 1µl 10µl primer and 2µl PCR product. The plate was sealed with an adhesive cover and put into the thermocycler, for 40 cycles of 96°C for 20 seconds, 50°C for 20 seconds and 60°C for 4 minutes.

2.10.3 Post sequencing PCR Clean-Up

Materials

- 70% and 95% Molecular Grade ethanol (Sigma)
- 3M Sodium Acetate Buffer Solution (Sigma)
- Plate centrifuge (Allegra 21R, Beckman Coulter, High Wycombe, UK.)
- Sequencing Plate and adhesive foil lids (Beckman Coulter)
- Sample Loading Solution (SLS) (Beckman Coulter)
- CEQ 8000 Sequencing Instrument (Beckman Coulter)

Method

7mls 95% ethanol and 280µl 3M NaOAc were mixed in a reagent reservoir and 50µl was added to each well of the sequencing plate using a multichannel pipette. This was vortexed gently then left at room temperature for 45 minutes. The contents of each well were transferred into the Beckman sequencing plate, and the plate

centrifuged at 2,800rpm for 1 hour at 4°C. Following this the plate was inverted on paper tissues and spun upside down at 500rpm for 1 minute, to remove any residual ethanol. 150µl of cold 70% ethanol was added, the plate re-sealed, and spun again for 10 minutes at 2800rpm. Again the plate was inverted on some tissue as before, and the washing step repeated once more. After the final inverted spin, the plate was either loaded straight onto the sequencer, or was frozen at -30°C. For loading onto the sequencer 20µl of Sample Loading Solution (SLS) was added to each well and overlaid with light mineral oil. The program LFR-1 was used on the Beckman CEQ8000. Using a capillary temperature of 45°C, denaturing temperature of 90°C and injection voltage of 2.0KV for 60 seconds, separation voltage was 8.0KV for 104 minutes.

2.10.4 Assembly of Contigs and assigning of Allele, Sequence type and Clonal Complex

Method

Forward and reverse sequences were assembled, trimmed and any anomalies edited by verification against the original sequence traces, using the program Bioedit Sequence Alignment Editor (copyright © Tom Hall 1997, Dept of Microbiology, North Carolina State University). By interrogation of the campylobacter MLST database (www.mlst.net/campylobacter) checked sequences, were assigned allele numbers. Once the 7 alleles had been assigned the website was used to assign the Sequence Type and Clonal Complex for each isolate.

2.11 Sequencing Method 2-based on the Applied Biosystems 3700 Sequencer at Oxford University

2.11.1 PCR Clean-Up

Materials

- Polyethylene glycol (PEG) (Sigma)
- 95% ethanol (Sigma)
- Nuclease free water (Promega)

Method

60µl of PEG was added to each well and left at room temperature for 30 minutes (or overnight). The plate was then centrifuged for 1 hour at 2750rpm, and then centrifuged upside down at 500rpm for 1 minute on tissues to remove any residual PEG. 150µl of ice-cold ethanol was added to each well and the plate again centrifuged at 2750rpm for 10 minutes, and inverted as before. This wash step was repeated once more, then 5µl sterile water was added to each well and the plate stored at 4°C.

2.11.2 Sequencing PCR

Materials

- Big Dye® Terminator v3.1 sequencing reagent (Applied Biosystems)
- Primers (10µM) -7 forward and 7 reverse (Table 2.2) (MWG Biotech)
- 96 well Sequencing Plate and adhesive lid (AB gene, Epsom, UK)

- Gene Amp PCR system 9700 (Applied Biosystems).

Method

1µl of PCR product was added into each well of the sequencing plate, followed by 2µl 10µM primer and 2µl Big Dye. PCR cycling consisted of 24 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 2 minutes.

2.11.3 Post Sequencing Clean-Up

Materials

- 95% ethanol and 70% ethanol (Sigma)
- 3M Sodium Acetate Sterile Solution pH 4.6 (Anachem, Luton, UK)
- Plate Centrifuge (Eppendorf 5810R)
- Nuclease free water (Promega)
- Vortex (Genie-2, Scientific Instruments)
- 3700 DNA Analyser (Applied Biosystems)

Method

For each plate 7mls 95% ethanol and 280µl 3M NaOAc (pH 4.6) were mixed in a reagent reservoir and 52µl was added to each well of the sequencing plate. This was vortexed, left at room temperature for 45 minutes, then centrifuged at 2750rpm for 1 hour. The plate was then inverted onto tissue paper and centrifuged for 1 minute at 500rpm. This washing step was repeated, and the plate inverted again to remove any residual ethanol. Sterile water was added to resuspend the DNA and the plate loaded onto the sequencer.

2.11.4 Assembly of contigs and assigning alleles by use of Staden programs

Method

All sequences were analysed by using the Staden suite of computer programs encompassing the programs STARS, PREGAP and GAP4. STARS (Sequence Typing Analysis and Retrieval System) (written by Man-suen Chan, www.molbiol.ox.aC. uk). This allowed the automatic assembly of contigs into the same database and calling of the alleles automatically. The plates for sequencing were set up so that the first half (48 wells) of the 96 well plate contained all the forward reactions and the second half (48 wells) contained the same samples but with the reverse reactions in the equivalent position on the other side of the plate. A Microsoft Excel file was created displaying this set up, which was saved and imported into STARS along with the raw trace files from the sequencer.

The STARS program was able to recognise this file input format and used this to assemble the forward and reverse sequences into a contig. The program also cut the assembled contigs to the correct length for the allele, which were then automatically matched to those on the MLST database by a link to the *C. jejuni* MLST website. A list file was generated of each sample number and allele number. Any sequences, which did not display an allele type, were checked using PREGAP and GAP4, programs for manual sequence aligning and editing.

These results produced by STARS were imported into an excel file, along with the results from the other loci. When the 7 digit allelic profile was complete the data was

imported into the MLST website to assign the Sequence Types and Clonal complexes.

Table 2.1 MLST primers for PCR amplification

Locus	Forward	Reverse
aspA	**A9** AGTACTAATGATGCTTATCC	**A10** ATTTCATCAATTTGTTCTTTGC
glnA	**A1** TAGGAACTTGGCATCATATTACC	**A2** TTGGACGAGCTTCTACTGGC
gltA	**A1** GGGCTTGACTTCTACAGCTACTTG	**A2** CCAAATAAAGTTGTCTTGGACGG
glyA	**A1** GAGTTAGAGCGTCAATGTGAAGG	A2 AAACCTCTGGCAGTAAGGGC
*pgm_*	**A7** TACTAATAATATCTTAGTAGG	**A8** CACAACATTTTTCATTTCTTTTTC
*tkt_*	**A1** TTTAAGTGCTGATATGGTGC	**A4** CATAGCGTGTTCTCTGATACC
uncA	**A7** ATGGACTTAAGAATATTATGGC	**A2** GCTAAGCGGAGAATAAGGTGG

(A-amplification, odd numbers indicates forward primer, even number indicates reverse)

Table 2.2 MLST primers for Sequencing

Locus	Forward	Reverse
aspA	**S3** CCAACTGCAAGATGCTGTACC	**S6** TTCATTTGCGGTAATACCATC
glnA	**S3** CATGCAATCAATGAAGAAAC	**S6** CCAAAGCGCACCAATACCTG
gltA	S1 GTGGCTATCCTATAGAGTGGC	S6 CCAAAGCGCACCAATACCTG
glyA	**S3** AGCTAATCAAGGTGTTTATGCGG	**S4** AGGTGATTATCCGTTCCATCGC
*pgm_*	**S5** GGTTTTAGATGTGGCTCATG	**S2** TCCAGAATAGCGAAATAAGG
*tkt_*	**S5** GCTTAGCAGATATTTTAAGTG	**S4** ACTTCTTCACCCAAAGGTGCG
uncA	**S5** TGTTGCAATTGGTCAAAAGC	**S4** TGCCTCATCTAAATCACTAGC

(S-sequencing, odd numbers indicates forward primer, even number indicates reverse)

Chapter 3

Applicability of PFGE and AFLP for outbreak investigations and utility for strain detection and characterisation

Chapter 3

Applicability of PFGE and AFLP for outbreak investigations and utility for strain detection and characterisation

3.1 Introduction

Molecular fingerprinting techniques aim to characterise strains based on the distinct patterns produced when chromosomal DNA, plasmid DNA or PCR products are subjected to treatment with restriction endonucleases. Both Pulsed Field Gel Electrophoresis (PFGE) and single enzyme Amplified Fragment Length Polymorphism (AFLP) are based on the enzymatic digestion of the whole genome, rather than just specific genes. Subsequently a high level of information can be obtained at the genetic level providing an overview of the entire genome.

PFGE has been used extensively for campylobacter characterisation to inform epidemiological studies (Gibson *et al.,* 1995; Lorenz *et al.,* 1998; On & Harrington 2001; Sails *et al.,* 2003c; Yan *et al.,* 1991) and for this reason is considered the gold standard genotyping technique for campylobacter. Extensive use of the technique has been reported for identification of human outbreak isolates (Allerberger *et al.,* 2003; Champion *et al.,* 2002; Lehner *et al.,* 2000; Olsen *et al.,* 2001), where in some cases discrimination has been reported to be equal or superior to other genotyping techniques (Champion *et al.,* 2002; Sails *et al.,* 2003c). Also PFGE has been used in passage studies through chickens (Hanninen *et al.,* 1999), for investigations of the diversity of strains within different chicken flocks (Perko-Makela *et al.,* 2002) and within poultry for retail sale (Dickins *et al.,* 2002). Furthermore, the technique has been used extensively for investigations of campylobacters within other animals and birds including cattle and dairy herds (Nielsen 2002), black headed gulls (Broman *et*

al., 2002), swine (Cloak & Fratamico 2002; Hume *et al.,* 2002) and reindeer (Hanninen *et al.,* 2002).

Additionally, PFGE has been exploited as a useful technique for localisation of *C. jejuni* genetic markers (Kim *et al.,* 1992) to predict the genome size (Taylor *et al.,* 1992) and can be used as a reliable means for differentiating between, and within, species of *Campylobacter* (Smith *et al.,* 2000b; Yan *et al.,* 1991). Also the technique has been used to investigate patterns of antibiotic resistance within different populations and habitats (Hein *et al.,* 2003; Nawaz *et al.,* 2003a; Pedersen & Wedderkopp, 2003; Wu *et al.,* 2002).

Standardisation of results from PFGE has become easier with the introduction of computer software to analyse and compare fingerprint patterns. Software such as Bionumerics (Applied Maths) or GelCompar (Applied Maths) has facilitated the comparison of gel fingerprints enabling results to be compared across laboratories and databases to be compiled, ultimately allowing the recognition of defined types. With the additional standardisation of materials and methods then interlaboratory PFGE comparisons become more feasible. This has been achieved for campylobacter with the campynet protocol in Europe, where a standardised protocol was devised and tested in a number of laboratories. Also this has been achieved in the US with a network of laboratories adopting the Pulsenet PFGE protocol, with standardised materials and methods.

A large number of studies have been carried out using the newer technique of Amplified Fragment Length Polymorphism for typing campylobacters (Desai *et al.,*

2001; Duim *et al.,* 2001; Hanninen *et al.,* 2001; On & Harrington, 2000). There are two variations in the way in which this technique can be applied. Most commonly reported is the fluorescent AFLP technique whereby two restriction endonucleases are used to digest the DNA, adapters are ligated onto the ends of the fragments, these are amplified in a PCR reaction with the addition of fluorescent dye terminators, then the fragments are analysed on an automated sequencer (Janssen *et al.,* 1996). The second variation of single enzyme AFLP is simpler and uses one restriction endonuclease, the ligation of adapters, then conventional PCR to amplify the fragments, which are analysed on agarose gels (Gibson *et al.,* 1998).

At the start of this investigation single enzyme AFLP (SAFLP) was a relatively new genotyping method which had only been applied to a few organisms these including *Salmonella enterica* serovar Typhi (Nair *et al.,* 2000) and *Helicobacter pylori* (Gibson *et al.,* 1998). Very little had been reported on the applicability of SAFLP for use in real time outbreak investigations and the discriminatory power of the technique when compared to other genotyping techniques for campylobacter, in particular PFGE. Additionally neither PFGE nor SAFLP had been extended to show use as a potential detection system within a non-culture system, where other interfering factors may be present.

FTA filters (Section 2.4.5) are described as a convenient method for the archiving and storage of DNA prior to PCR methodologies. It was anticipated that these could not be used in conjunction with the PFGE due to the necessity for the DNA to be contained within agarose blocks, however the FTA filters used with the SAFLP technique would seem feasible. FTA filters have been used successfully in

conjunction with PCR as a detection method for *Listeria monocytogenes, Shigella flexneri*, and *S. typhimurium* (Lampel *et al.*, 2000) and used for the collection of DNA from blood (Devost & Choy, 2000). However the filters have not been described in conjunction with any molecular fingerprinting techniques.

The overall aim of this thesis was to establish timely discriminatory strain fingerprinting for early recognition of case clusters of *C. jejuni* infection, therefore the aims of this chapter were two fold. In the first instance it was proposed to utilise and compare the methods, PFGE and SAFLP for molecular fingerprinting of campylobacter for characterisation of outbreak isolates (In collaboration with Olivia Champion, London School of Hygiene and Tropical Medicine, Keppel Street, London). Secondly, the potential for applying these techniques in possible detection and characterisation of campylobacter specific strains directly from non-cultured samples and by the use of archived DNA on FTA filters was investigated.

3.2 Materials and Methods

3.2.1 Pulsed Field Gel Electrophoresis (PFGE)

PFGE was carried out by the Laboratory of Enteric Pathogens campylobacter PFGE protocol (as described in section 2.7.1). The restriction endonuclease *Sma*I was used throughout, and all analyses were carried out by the use of Bionumerics (as described in Section 2.8).

3.2.2 Single Enzyme Amplified Fragment Length Polymorphism (SAFLP)

SAFLP was carried out according to the method of Gibson *et al* (Gibson *et al.,* 1998) using the restriction endonuclease *Hin*dIII (as described in section 2.7.2). Analysis was performed using Bionumerics as described for PFGE.

3.2.3 Adapted Single Enzyme Amplified Fragment Length Polymorphism (SAFLP) technique

The method was modified slightly to the above by the use of a combined digestion and ligation step carried out at 37°C for 1 hour (as described in section 2.7.3).

3.2.4 Limits of Detection for the SAFLP Technique.

For PFGE, the process of embedding campylobacter DNA within agarose blocks to protect the DNA during the subsequent lysing and washing steps was critical. Therefore it could be anticipated that the use of PFGE on any type of sample material other than cultured cells would be unsuccessful. For this reason only SAFLP was investigated for application to both culture and non-cultured samples.

All the analysis of campylobacter by SAFLP to date had been carried out on a standard concentration (0.266μg/μl) of pure isolated DNA. The application of SAFLP to a range of different concentrations of DNA was explored in order to determine the limits of detection for the technique. Triplicate dilutions of RNase treated DNA (as described in Section 2.4.6) were prepared from 100 to $1x10^{-7}$ μg/μl and tested with the SAFLP procedure.

3.2.4 SAFLP analysis of archived DNA stored on FTA filters

DNA had previously been archived on FTA filters (Section 2.4.5). One FTA punch of archived campylobacter DNA was added into the first SAFLP digestion/ligation mastermix and the volume made up to 25μl with distilled water. The remainder of the SAFLP process was carried out as described previously. This process was carried out also after six months, using "punches" from the same archived DNA cards to test the stability of the DNA stored on cards over time.

3.3 Results

3.3.1 Comparison of PFGE and SAFLP Techniques for Outbreak Investigation (In collaboration with Olivia Champion section 3.3.1)

An outbreak investigated by the Campylobacter Reference Unit was used to compare the discriminatory power of the techniques PFGE and SAFLP for use in identifying outbreak isolates.

(i) Campylobacter Outbreak

Epidemiology

In September 2000 a local increase in *C. jejuni* isolation rates occurred in South Wales involving three separate groups of patients within a 35-mile radius. These included (i) seven residents of a nursing home (HS50 PT6) (Incident 1), (ii) one resident of a housing estate (HS23 PT1) (Incident 2) and (iii) one cot death (HS UT PT RDNC) (Incident 3). Thirteen sporadic isolates received into the reference laboratory during the same time period were also analysed.

PFGE and SAFLP Data

PFGE was applied and the resulting 6 to 8 fragments per strain ranged in size from 45-3600 kb. When analysed using Bionumerics, PFGE clustered all seven of the HS50 PT6 isolates from the nursing home as 100% identical. Four sporadic isolates (all HS50 PT6) strains also clustered within this profile. The other potential outbreak cases (housing estate and cot death) both involved single isolates with unique profiles distinct from each other and all of the other isolates tested (Figure 3.1).

By SAFLP 8-10 fragments were generated ranging in size from 400-1000bp. SAFLP clustered the HS50 PT6 outbreak strains from the nursing home and four sporadic HS50 PT6 strains as 100% identical. The other cases (housing estate and cot death) showed distinct profiles to the isolates from the nursing home although four sporadic strains with different serotype and phagetype clustered with the profile of the cot death isolate with 100% homology (Figure 3.2).

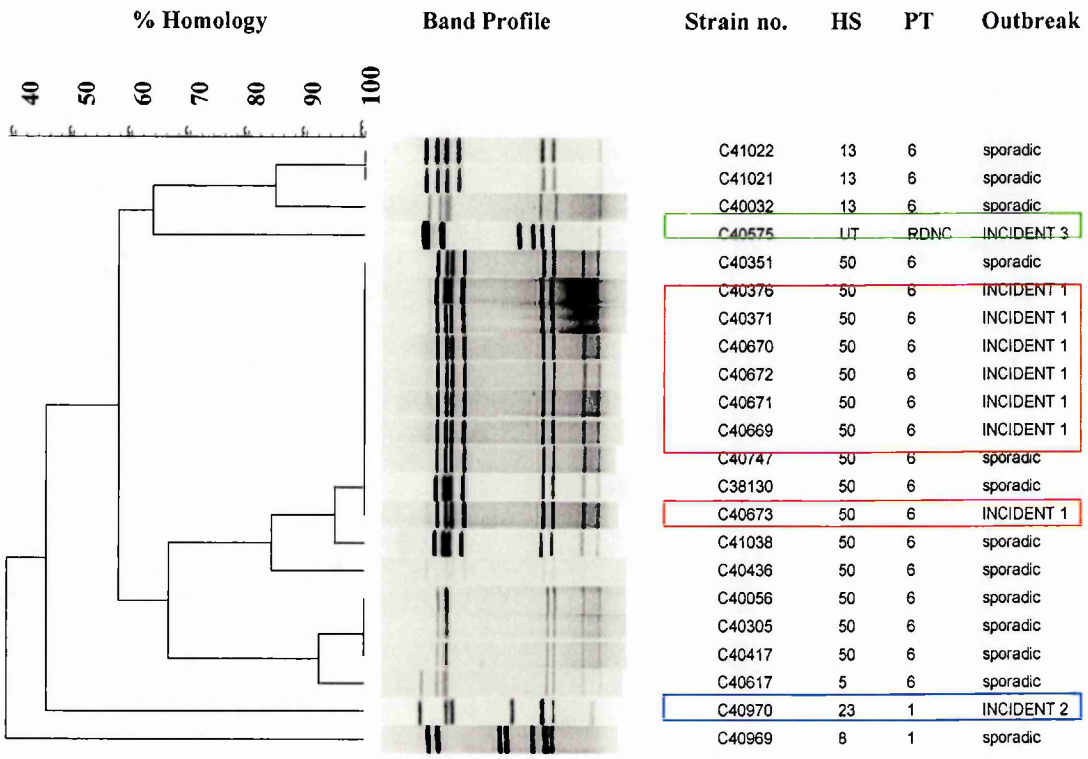


Figure 3.1 Dendrogram of PFGE patterns digested with *Sma*I. Cluster analysis was performed with Bionumerics (Applied Maths, Kortrijk, Belgium) using Dice correlation coefficient and UPGMA clustering algorithm. (Red-Incident 1, Blue-Incident 2 and Green-Incident 3). Between 6 and 8 fragments were produced. The seven isolates from Incident 1 (HS 50 PT6) clustered as 100% homologous, four sporadic isolates also clustered with this profile. Incidents 2 and 3 involved single isolates with unique profiles.

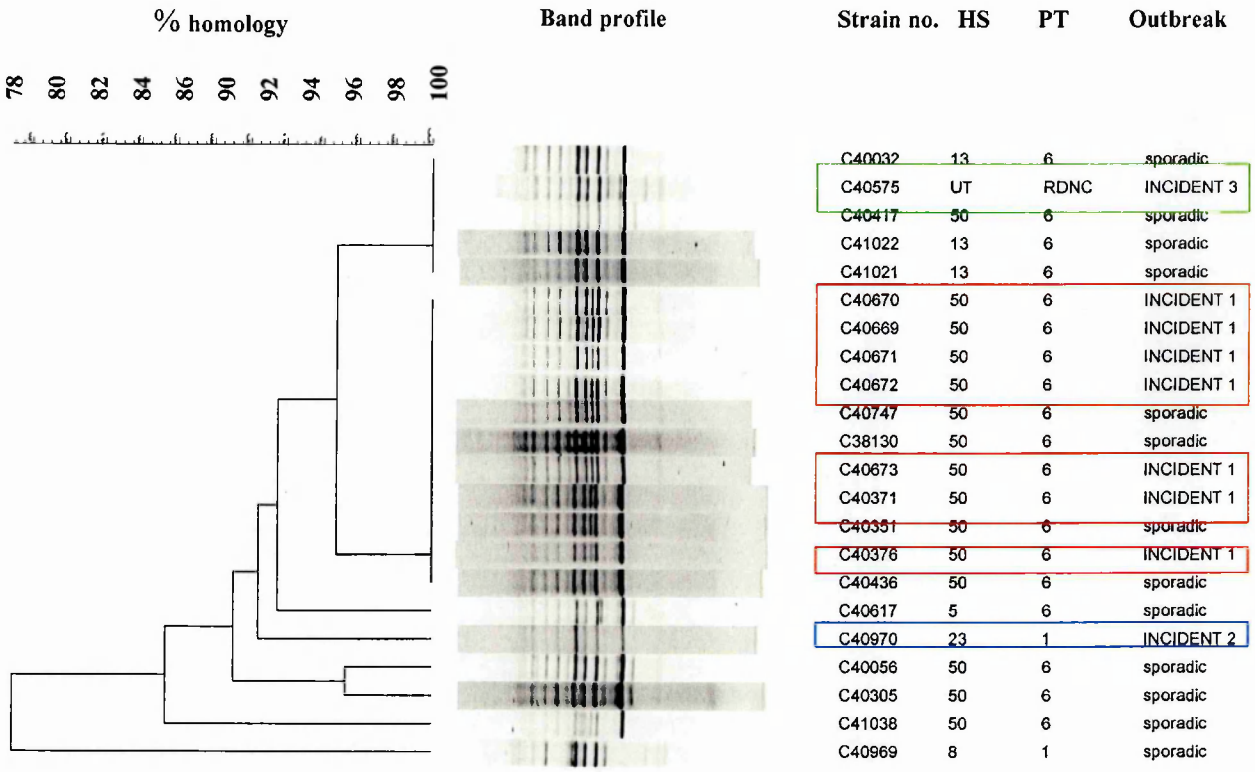


Figure 3.2 Dendrogram of SAFLP digested with *Hin*dIII. Cluster analysis was performed with Bionumerics (Applied Maths, Kortrijk, Belgium) using Dice correlation coefficient and UPGMA clustering algorithm. (Red-Incident 1, Blue-Incident 2 and Green-Incident 3). Eight to ten fragments were produced. All Incident 1 (HS50, PT6) isolates clustered with 100% homology including four sporadic isolates (HS50, PT6). Both Incident 2 and 3 isolates showed unique profiles.

3.3.2 (i) Determination of the Limits of Detection for the SAFLP Technique.

Successful fingerprints were obtained by the combined digestion/ligation step carried out at 37°C for 1 hour. No variation in results was seen and the method provided a quicker alternative.

The mean range of DNA, which could be successfully detected by SAFLP, was between 1μg/μl and 0.1μg/μl. (Figure 3.3). Both lanes containing higher concentrations of DNA (Lane 1 containing 100μg/μl and Lane 2 containing 10μg/μl) failed to work correctly. Likewise the lanes containing the concentration of DNA below the concentration of 0.1μg/μl failed to work. This result was consistent with the DNA starting concentration as described by Gibson *et al* (1998) of 0.266μg/μl, that the SAFLP technique can only accommodate a very defined range of possible starting concentrations. For these reasons further investigation of the SAFLP technique directly from culture, or other potential starting materials was not attempted.

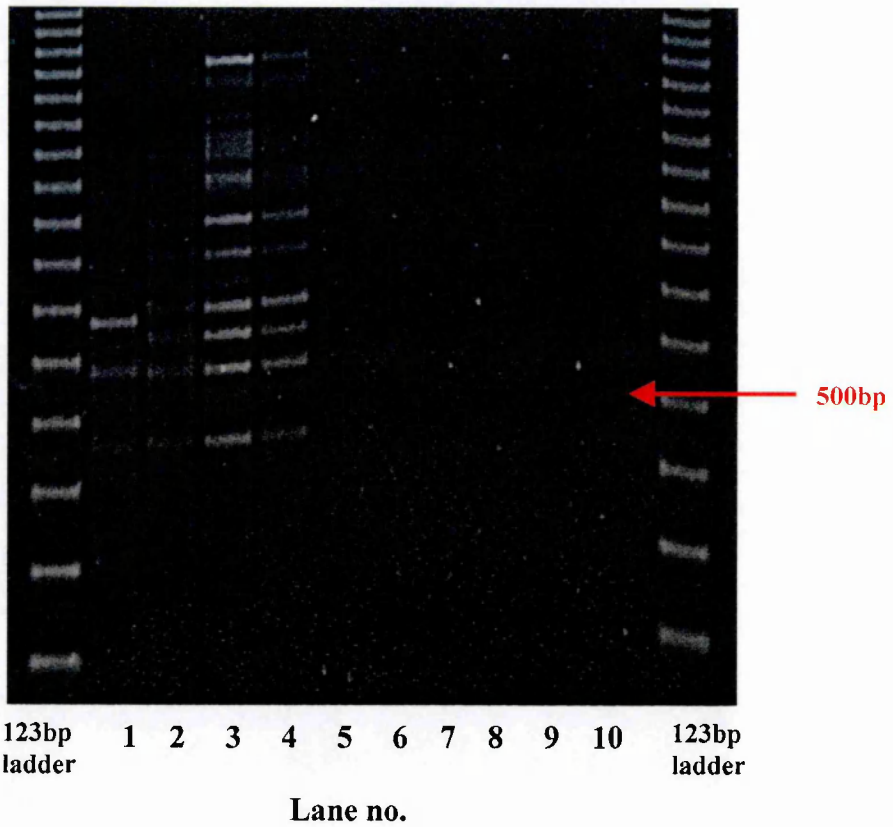


Figure 3.3 SAFLP fingerprint produced by serial dilutions of DNA from 100-1x10^{-7} µg/µl. Lanes 1 (100 µg/µl) and 2 (10µg/µl) showing incomplete resolution of banding patterns due to too much DNA within the reaction. Clear banding patterns have been produced for lanes 3 and 4 containing 1µg/µl and 0.1µg/µl DNA respectively. Lanes 5-10 show no clear banding patterns due to insufficient DNA within the reaction (0.01-1x10^{-7} µg/µl).

3.3.2 (ii) Use of archived DNA on FTA filters

Standard SAFLP was carried out on extracted DNA diluted to the required concentration of 0.266µg/µl (Figure 3.4). Approximately 8-10 bands were produced ranging in size from 500bp to 1845bp. Clustering of strains was clearly visible with the profiles in lanes 1,5,7,8,9 and 10 showing similar profiles. Lane 2 showed a similar fingerprint but with one band difference. Lanes 3 and 4 showed identical fingerprints, which were distinct from the other profiles. Likewise lane 6 was also distinct from the other profiles with absent smaller bands.

Figure 3.5 shows the SAFLP fingerprint from DNA stored on FTA paper. The profile resembled the fingerprint in Figure 3.4 with consistent relationships between the different profiles however there were subtle changes to the banding patterns. Extra small bands appeared at the bottom of the gel after storage of the DNA on the FTA paper. It was still possible to distinguish the fingerprints however the patterns were different. The bands seen at the bottom of the gel were between 200-400 base pairs, considerably smaller than the smallest fragments seen on the original fingerprint.

From gel picture Figure 3.5 it was still clear that the profiles in lanes 7-10 were identical, each showing the same profile. With the presence of the smaller bands it was more difficult to distinguish the types as some of the smaller bands did not resolve correctly due to their size. This would cause problems if the gel image were to be analysed in Bionumerics, as the presence of these extra bands would have to be included in the analysis.

Following six months of storage of the DNA on FTA cards the SAFLP procedure was repeated. Substantial DNA degradation had occurred and fingerprints produced did not resemble those produced previously, a number of fingerprints did not resolve at all (results not shown).

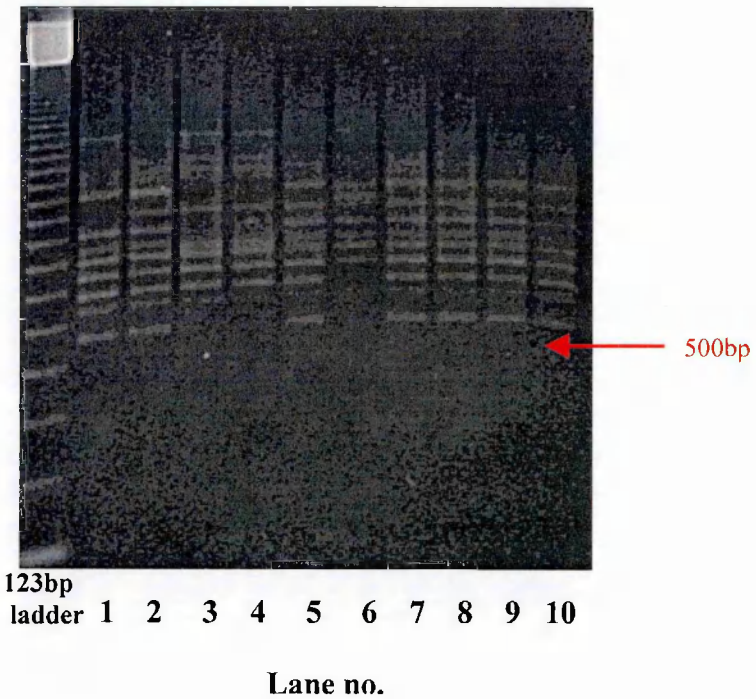


Figure 3.4 SAFLP gel image from standardised extracted DNA (0.266µg/µl). 8-10 bands were produced from 500bp to 1845bp. The profiles in lanes 1,5,7,8,9 and 10 showed similar profiles. Lane 2 showed a similar fingerprint but with one band difference. Lanes 3 and 4 showed identical fingerprints, which were distinct from the other profiles. Lane 6 was also distinct from the other profiles.

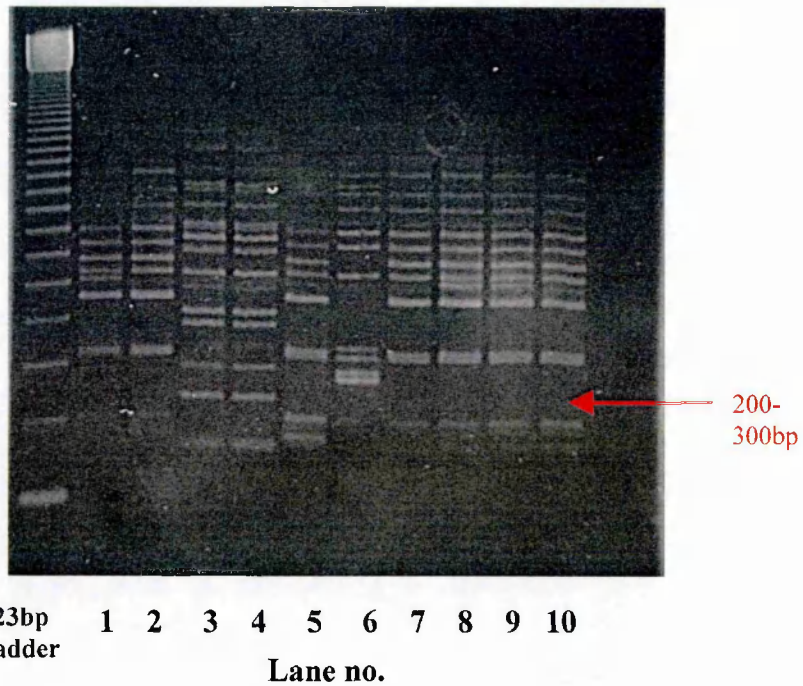


Figure 3.5 SAFLP gel image from DNA stored on FTA cards. 10-15 bands were produced from 200bp to 1845bp. Profile resembled the fingerprint in Figure 3.4 with consistent relationships between the different profiles. The profiles in lanes 1,5,7,8,9 and 10 showed similar profiles. Lane 2 showed a similar fingerprint but with two bands difference. Lanes 3 and 4 showed identical fingerprints, which were distinct from the other profiles. Lane 6 was also distinct. Smaller bands (200-300bp) are present on the gel due to fragmentation of the DNA.

3.4.Discussion

3.4.1 Comparison of PFGE and SAFLP for identification of outbreaks

The fundamental attribute of a molecular fingerprinting method is its capacity to differentiate between epidemiologically-related and unrelated strains. It is imperative for clinical and epidemiological studies that *C. jejuni* strains are accurately differentiated. Both SAFLP and PFGE were successful in identifying the different outbreak strains from each other as well as the sporadic strains. Both these techniques grouped some of the sporadic strains with the outbreak strains, demonstrating the importance of typing both outbreak and sporadic strains in order to estimate the population prevalence of specific types. Since *C. jejuni* is not highly clonal different typing methods may group isolates differently and generate different clusters. However within a single clone all isolates should be the same with any combination of typing techniques.

The discriminatory power of PFGE fingerprinting for campylobacter is high, although the technique is time consuming. The process requires multiple labour intensive steps whereby the DNA is prepared within agarose blocks, subsequently washed and lysed. The electrophoresis stage is also lengthy with 24-48 hours running time required for the migration of DNA fragments. Also the equipment for electrophoresis is highly specialised and likely to be restricted to research and reference laboratories.

Although PFGE is considered the gold standard for molecular typing of campylobacter isolates one of its potential drawbacks is the necessity to isolate intact chromosomal DNA within the agarose blocks. These agarose blocks protect the

DNA from mechanical shearing. The slow processes of subsequent lysis of cells with a detergent and degradation of the proteins with a protease means the DNA is potentially exposed to nucleases for a relatively long time. This may explain some of the problems with the technique for certain isolates where DNA degradation occurs due to certain strain characteristics. These problems can be avoided by additional steps, such as treatment of the cells with formalin (Liesegang & Tschape, 2002), by the use of a different or complete elimination of proteinase (Turabelidze *et al.,* 2000), by use of different running buffers (Koort *et al.,* 2002; Romling & Tummler, 2000) or by the addition of a free radical scavenger such as thiourea during the electrophoresis stage (Liesegang & Tschape, 2002); all add to the complexity of the procedure. PFGE is not a failsafe procedure, however it remains one of the few methods in which very large DNA fragments of the whole genome can be analysed electrophoretically.

In other methods for DNA extraction such the Isoquick DNA extraction method rapid inactivation of nucleases occurs by the addition of chaotrophic salts, but the DNA isolated is not completely intact and tends to be sheared to fragments of approximately 50kb making it unsuitable for PFGE analysis. For these reasons PFGE cannot be applied directly to potential sources of campylobacter without a primary culturing step.

Many methods for PFGE exist for campylobacter, the one used within this study was the PFGE protocol from the Laboratory of Enteric Pathogens, which was chosen due to its reproducibility in previous studies. However two other common protocols in use for campylobacter are the Campynet protocol and the Pulsenet protocol. Both

methods were devised to reduce some of the problems relating to interlaboratory standardisation by the introduction of universal reagents and electrophoretic parameters. However these three methods produce comparable patterns and have much in common.

Many attempts have been made to improve the technique of PFGE in order to speed up the process, without compromising the quality of results. A rapid method for PFGE was developed by the network of PulseNET (Ribot *et al.*, 2001) laboratories in the USA. Various concentrations of buffers reagents, reaction time and electrophoretic parameters were evaluated in an effort to devise a protocol that was rapid, simple and robust. The revised protocol allowed for PFGE to be completed within 24 hours rather than the 2-3 days required for the conventional PFGE protocols, however the technique was still very time consuming with a great deal of time spent at the bench. Similar rapid methods have also been reported for other organisms, for example in typing of vancomycin-resistant enterococci (Turabelidze *et al.*, 2000), *S. enterica* and *Escherichia coli* (Liesegang & Tschape, 2002), however none of these describe a protocol achievable in less than 24 hours.

One of the main issues surrounding the use of molecular fingerprinting techniques is the issue of reproducibility and comparability between different laboratories. The Pulsenet method was designed for use in a network of laboratories within the US, the primary aim of which was to standardise PFGE and enable gel fingerprints to be compared. The network of laboratories achieved this goal and PFGE is being carried out in multiple locations with comparable results. However this is at the expense of strict adherence to standardised protocols and highly time consuming normalisation

and interpretation of PFGE fingerprints. A similar system was established in Europe known as the Campynet project, where the standardised method was implemented, however as in the case of Pulsenet, it was only a success by strict adherence to the set methods at every step in the procedure. Likewise, similar strategies have been devised with other organisms for example the Salm-gene project for Salmonella (Peters *et al.*, 2003), is a European wide collaboration where harmonisation of PFGE has achieved up to 95% similarity between PFGE profiles between laboratories.

The process of SAFLP is less time consuming and requires standard laboratory equipment, a thermocycler and agarose gel electrophoresis tanks. SAFLP requires a standard amount of relatively pure DNA, which does not have to be completely intact as for the case of PFGE allowing for a variety of DNA extraction procedures to be applied. The sizes of the genomic fragments, which are analysed by SAFLP, are within the range of 500-2000bp. This makes the SAFLP method more tolerant of poorer quality DNA than PFGE and even allows SAFLP to be successful on slightly fragmented DNA. The procedure for SAFLP in total hours is less than the complete procedure for PFGE and also requires substantially less hands on time. The SAFLP procedure can be completed within 8 hours if the modified method is used, allowing one hour for a combined digestion and ligation step, three hours for the PCR amplification stage, and four hours for the electrophoresis and ethidium bromide staining. By comparison, the PFGE technique at its fastest, using the PulseNET technique takes 24 hours with considerably more hands on time.

For both techniques additional time for the fingerprint analysis is required. Despite the use of computer software such as Bionumerics for gel analyses, the process is

still time consuming especially for gel normalisation, crucial for comparisons against existing gels in a database. This step is vital for standardisation and to enable the technique to be comparable across laboratories. In order for SAFLP to be used as a standardised typing technique as described for Pulsenet (Ribot *et al.,* 2001), then designated type strains would have to be included into every gel to allow for standardisation and every step of the process this would have to be done in accordance with well defined procedures. A standard nomenclature would also need to be established whereby a set fingerprint could be assigned a type based on the number of bands present. Both Pulsenet and the Salm-gene projects (Peters *et al.,* 2003) have dealt with this by assigning fingerprint designations at a central laboratory. This offers the advantage that all the gels will be analysed by the same techniques, however it adds a time penalty for interlaboratory comparisons.

The other problem compounding PFGE is the limitation on the starting material. The whole technique is based upon the availability of sufficient bacterial growth in order to prepare the initial saturated suspension of cells, and using too few or too many initial cells compromises results. Likewise, it is not a technique applicable for the detection and identification of campylobacters directly from isolated DNA if culture was unavailable. The length of the process, and the requirement for specialised equipment, make it unsuitable as a rapid method for detection in front line laboratories. Moreover, further development of PFGE is limited by the lack of restriction enzymes available due to the low G+C content of the campylobacter genome The most commonly used restriction enzymes are *Sma*I and *Kpn*I. *Sma*I targets high GC sequence sites (CCC↓GGG) therefore for campylobacter with a low GC genome, this enzyme cuts infrequently. *Kpn*I is similar in that it cuts high GC

sequence types (GGTAC↓C). Conversely, for SAFLP, the frequent cutter *Hind*III (A↓AGCTT) is used. The SAFLP process was not designed for studies on campylobacter therefore the choice of restriction enzymes could be altered to improve discrimination. For FAFLP described for campylobacter other enzymes are used including *Hin*dIII and *Hha*I (Desai *et al.,* 2001; Duim *et al.,* 2001) and *Bgl*II and *Csp*6I (On & Harrington 2000). Extensive evaluations would need to be carried out for alternative enzymes to determine the number of fragments produced and whether they would be of a suitable size and quantity to discriminate strains effectively by conventional agarose gel electrophoresis.

PFGE is applicable and has been described for other species of *Campylobacter.* The technique can be used for distinguishing *C. jejuni and C. coli* (Cloak & Fratamico, 2002). It has also been applied to *C. hyointestinalis* (Gorkiewicz *et al.,* 2002; Hanninen *et al.,* 2002), *C. consiscus* (Matsheka *et al.,* 2002), *C. fetus* (Fujita *et al.,* 1995; Morooka *et al.,* 1996; Vargas *et al.,* 2003) as well as urease positive thermophillic *Campylobacters* (UPTC) (Matsuda *et al.,* 2002). However to date no such work has been described for SAFLP using other species of *Campylobacter.*

The finding that SAFLP has equal discriminatory ability to PFGE in the identification of outbreak isolates is in line with another study comparing the two techniques for nosocomial pathogens (D'Agata *et al.,* 2001). This study reported the two techniques to have equal discriminatory power for the organisms *A.baumani* and *P.aeruginosa* however PFGE remained the method of choice for *Enterococcus faecium.*

This study was designed to evaluate the utility of PFGE and SAFLP for outbreak investigations and it is concluded that both these methods have approximately equal utility. The discriminating potential of PFGE is excellent however the disadvantages of the technique lie in the lengthy preparation process for the DNA samples and its dependence on specialised and expensive equipment (Frost, 2001). For investigation of future outbreaks SAFLP with its ease of use, rapid turnaround and availability of equipment make it an attractive alternative to PFGE in a local context. Future investigations would need to be carried out to establish the merits of the procedure for interlaboratory comparisons.

3.4.2. Investigation of the Utility of SAFLP as a Method for Detection

The second aim of this chapter was to investigate the applicability of the use of molecular fingerprinting techniques for timely detection of specific strains of campylobacter to enable a public health response to establish the source of infection. A method for rapid strain characterisation has to be sufficiently robust, sensitive, highly specific and ideally identify strains to a recognised type by known nomenclature. Moreover, it has to be adapted to non-culture samples in order to achieve rapid results.

Despite the success of PFGE as a method for discriminating strains of *Campylobacter*, described in this chapter and elsewhere, (Champion *et al.,* 2002; Ribot *et al.,* 2001; Sails *et al.,* 2003c) PFGE was considered unsuitable as a technique for rapid characterisation of strains in the context of this project for two main reasons. Firstly the PFGE technique is time consuming and tedious, especially if the data analysis requires strict normalisation to aid comparability between laboratories. Secondly, the starting material required restricts the technique to cultured samples and the technique cannot be applied to food or environmental samples without a prior culturing stage.

To be applicable for a method of detection for specific campylobacter strains there are two main requirements, neither of which is met by PFGE. In the first instance different strains of campylobacter have to be recognised by a defined type with strict nomenclature in order to aid epidemiological studies. PFGE although applicable for the identification of outbreak isolates currently has no recognised nomenclature for the different fingerprints produced. With the absence of recognised type strains then

the technique has no value for specific identification of campylobacter strains. If standardised procedures were in place, and a recognised nomenclature could be established then PFGE may be more applicable for detection. Standardising of methods and set type strains are crucial for the successful implementation of any molecular method as a practical epidemiological tool. For PFGE this would require a great deal of effort and organisation, which is perhaps not realistic when there are other newer typing techniques coming into development for campylobacter.

SAFLP was determined to have equal discriminatory ability for distinguishing outbreak strains of campylobacter; the technique could be applied to extracted DNA and was quicker to perform than PFGE. For these reasons SAFLP was a more viable alternative to test as a possible strategy for rapid strain detection and characterisation. Initially it was attempted to test the SAFLP technique on non-standardised quantities of DNA to determine the limits of detection and utility of the system. This was especially important, as a method for rapid epidemiological characterisation would have to be sufficiently adaptable to the varying amounts of campylobacter DNA found within the food and environmental specimens likely to be analysed which would also speed the process up. A system whereby defined quantities of DNA were required would be impractical for non-culture strain characterisation.

It was possible to reduce the overall time taken for the SAFLP process by use of a combined ligation and digestion mix with a shorter incubation time of 1 hour. This reduced the overall time required for the technique by 3 hours, and made the whole process achievable within one working day. Further reduction in time would not be possible due to the constraints of the PCR step and the slow electrophoresis due to

the low voltage and high concentration of agarose gel required in order to resolve the small fragments (Wilson *et al.*, 1987a).

Varying concentrations of DNA were tested in the SAFLP procedure to determine the limits of starting material for the technique and whether it could be used for detection purposes. By using a range of serially diluted DNA it was only possible to produce good quality SAFLP fingerprints within the DNA range of 0.1-1μg/μl. This finding was consistent with the recommended amount of starting material as described by Gibson *et al* (1998). The finding that SAFLP is only effective within defined ranges of starting material makes that process unsuitable for a method that could be applied for detection and characterisation of campylobacter strains within potential sources.

The importance of detecting small quantities of campylobacter DNA is imperative for the design of a sensitive detection method. The only way in which this method could be applied to environmental or food samples would be by enrichment of the sample followed by DNA extraction, quantification of the DNA, followed by the SAFLP technique. Alternatively, there are new PCR reaction kits available, which enable the homogenous amplification of genomic DNA without specific primers, such as the Genomiphi kit (Amersham Biosciences). These approaches could offer a strategy whereby more starting material could be obtained, however they add time onto the process (by 24 hours for enrichment or 12 hours for the Genomiphi amplification) which is beyond the limits of timely detection and characterisation of campylobacters.

Use of the FTA cards was attempted in order to speed up the process of SAFLP by reducing the time for the DNA extraction step and for the expected benefits of the storage facility enabling comparisons over time. The use of FTA cards in conjunction with the SAFLP technique was not entirely successful. The appearance of extra small bands within the SAFLP profile suggests that the DNA was becoming fragmented on storage resulting in the small extra bands not seen in the profile obtained from standard extracted DNA. It would seem likely that the fragmentation was occurring as result of the fixing process and thereafter as a result of the fixed DNA being stored, accounting for the loss in SAFLP fragments.

A further problem with the cards is the unknown quantity of DNA, which is initially applied onto the card. Although a standard bacterial suspension was added it would be difficult to quantify how much resulting DNA is present and available to be modified during the SAFLP procedure per punch. Also the fact that the DNA remains immobilised to the card during the SAFLP process suggests that the amount of DNA available to react is only going to be the DNA which is released from the matrix through the digestion and ligation process and free in the solution. This is especially important when a critical DNA concentration range is required for SAFLP to work successfully.

With other standard PCR techniques the cards have been reported as successful. For example in blood collection for PCR analysis (Devost & Choy, 2000) and also in the detection of the food borne pathogens *Shigella flexneri*, *Salmonella enterica* serotype Typhimurium and *Listeria monocytogenes* by PCR (Lampel *et al.*, 2000) however no application of the cards with SAFLP procedures has been described.

The resulting extra bands seen suggest that this would not be a suitable alternative for DNA extraction to use in conjunction with the SAFLP procedure. The main problem being the lack of comparability with existing profiles, and rearrangement of the banding patterns. Standardisation with existing profile databases would be impossible and even if a degree of standardisation could be achieved there would be no indication that the fingerprints produced would be stable over time. Further storage may result in more DNA degradation therefore further divergence of the banding patterns.

This method of using DNA stored on cards would be highly desirable as an alternative to manual extraction procedures, and would significantly aid epidemiological studies if used in conjunction with applicable rapid techniques. This would allow for convenient purification of DNA whenever required and the ability to store DNA at room temperature without occupying freezer space, as well as the availability of large reserves of DNA for analysis. However with the limitations on the amount of DNA, which can be analysed for SAFLP, the process is not viable without substantial further development.

Despite PFGE and SAFLP being successful methods for the recognition of epidemiological groupings of campylobacter they are unable to provide the timely fingerprinting of isolates for robust early detection and case cluster recognition. At the time at which these investigations were being carried out there was a great deal of development in the area of sequence typing, with real time sequence based detection systems such as the Taqman or Lightcycler becoming more widely available within

laboratories. Also the cost of full DNA sequencing is becoming less expensive with realistic and reliable sequencing being possible with relative ease by the use of automated capillary sequencers.

For these reasons the remainder of this thesis describes the development of further molecular sequence based techniques which have the applicability for application to potential sources of infection such as food or environmental samples and direct detection of specific strains.

Chapter 4

Development and applicability of a real time Taqman assay for identification of *C.jejuni* and *C.coli*

Chapter 4

Development and Applicability of a Real Time Taqman Assay for Identification of *C. jejuni* and *C. coli*

4.1 Background

C. jejuni and *C. coli* speciation is rarely undertaken in the majority of front line laboratories despite accurate speciation being important for clinical and epidemiological purposes. Differences in antibiotic resistance exist, especially in relation to erythromycin and ciprofloxacin, where resistance is more common in *C. coli* (Thwaites & Frost, 1997). Furthermore, a population based sentinel surveillance scheme from England and Wales has demonstrated the higher risk of obtaining an infection from a particular species by the exposure to certain food products. These hypotheses suggest differences between the two species with respect to different niches and transmission routes (Gillespie *et al.,* 2002).

Differentiation between *C. jejuni* and *C. coli* has traditionally depended upon the presence of hippuricase in *C. jejuni* and the absence in *C. coli* either by the phenotypic hippurate hydrolysis test (Bolton *et al.,* 1984a) or the detection of the *hip*O gene by PCR (Totten *et al.,* 1987). However, this has been determined to be an unreliable indicator of species due to the presence of atypical hippurate negative *C. jejuni* strains (Totten *et al.,* 1987). As a result many alternative PCR targets have been described based on genes or parts of genes specific to either species that have superseded the use of the hippurate for differentiation (Table 1.2).

A comparison of eleven *C. jejuni/coli* conventional PCRs was carried out by On and Jordan (2003) where considerable variation was observed in performance and no test was found to be 100% accurate. This was attributed to both biological factors including the natural diversity of campylobacter and also physical factors, relating to the technicalities of the PCR reaction. The authors' conclusion was that the use of more than one PCR assay was most reliable, but the Vandamme PCR (1997) was cited as the best with consistent results based on tests with 111 type, reference and field isolates. The authors described reproducibility problems with all the PCRs tested, suggesting that the assays may be optimal in the hands of the developers, although not optimal when slight variations in equipment or reagents were used. Additionally, many of the assays had been evaluated using DNA isolated from cultures and had not been extended to show detection from potential campylobacter sources where PCR interfering factors could be present.

Alternative approaches have been developed involving a PCR stage before or after an additional technique to increase the sensitivity. For example, enrichment for 24 hours followed by isolation and concentration by surface adhesion onto a polycarbonate membrane, and detection using PCR (SA-PCR), was used by Cloak *et al* (2001). PCR technologies have also been used in conjunction with ELISA (Enzyme-Linked Immunosorbent Assay) techniques, (PCR ELISA) these techniques have been used successfully for the identification of *Campylobacter* species. Using an initial polymerase chain reaction assay then subsequent detection based on a solution-hybridisation colorimetric end point detection format (PCR ELISA) provides a rapid alternative to gel based detection (Bolton *et al.,* 2002; Kulkarni *et al.,* 2002; Lawson *et al.,* 1999; Metherell *et al.,* 1999; Sails *et al.,* 2001). The

increased sensitivity achieved through the PCR ELISA approach allowed the extension of some of these assays to a food based system, for example that developed by Bolton *et al* (2002). Also commercial PCR kits have become available, such as the campylobacter BAX® PCR which was designed to identify *C. jejuni* and *C. coli* (Englen & Fedorka-Cray, 2002).

At the start of this investigation methods used for speciation of *C. jejuni* and *C. coli* within the CRU laboratory included a phenotypic hippurate hydrolysis test, biotyping (Bolton *et al.,* 1984a) and a multiplex PCR described by Denis *et al* (1999). This conventional PCR involved the amplification of a genus 16S rRNA sequence and sequences of two genes specific for either species, the *mapA* gene from *C. jejuni* and the *ceuE* gene from *C. coli.* The species specificity of these methods was confirmed by comparison with the serotyping, phagetyping and biotyping results. Despite the success achieved with using this conventional methodology the process was time consuming with consumable costs of £1.80 per isolate (including biotyping costs) The PCR involved a three-hour PCR amplification step plus time for electrophoresis, and was therefore not suitable for high throughput studies.

Real-time PCR assays based on platforms such as the Taqman (Applied Biosystems) show the greatest promise for large scale testing, and are favoured due to increased sensitivity, reproducibility and quick turnaround time (Guiver *et al.,* 2000). Assays have been described in increasing numbers over the last three years based on real time platforms across a wide range of disciplines, including bacteriology, virology, mycology and medical genetics. Although a large range of assays have been described many of them have been focused to laboratories where there is a

requirement for high throughput identification assays or quantification studies, for example in the quantification of HIV proviral load (Zhao *et al.*, 2002). Assays have been described for identification and quantification of *C. jejuni* (Nogva *et al.*, 2000a; Sails *et al.*, 2003a). There are also assays based on the Roche Lightcycler (Cheng & Griffiths, 2003; Logan *et al.*, 2001) for identification of *Campylobacter* species based on melting peak profiles, but at the time of writing (1/10/01) no Taqman assay for the speciation of *C. jejuni* or *C. coli* had been described.

DNA of sufficient yield and quality is vital for efficient performance in downstream applications and is an area often overlooked when designing PCR assays. Different extraction techniques may result in variable sensitivity of PCR detection and also different methods for extraction represent substantial differences in time required and cost. Many methods for DNA extraction are currently available including walk away robotic systems (such as the Roche MagNApure or Qiagen Biorobot), specific DNA extraction kits, traditional methods based on guanidinium thiocyanate (GuSCN) (Boom *et al.*, 1990) or CTAB (Wilson *et al.*, 1987b), and more novel methods by use of the autoclave method (Simmon *et al.*, 2004). DNA extraction needs to be tailored for a specific purpose, depending on the number of samples, quality and quantity of DNA required and cost. It would have been impossible to evaluate all types of DNA extraction methods available, therefore the MagNApure was tested due to its reported accuracy and rapidity and compared with the crude cell lysate preparation method as this was known to work in conventional PCR assays. Additionally methods for the isolation of bacterial DNA directly from food samples were tested in conjunction with this assay. This also included the use of the MagNApure, a method based on paramagnetic beads, and a novel DNA extraction reagent.

4.1.1 Aims of the Investigation

The aim here was to develop and evaluate a technique based on the Taqman real time PCR system, which could successfully and rapidly detect and identify the species *C. jejuni* and *C. coli*. There was a need for a sensitive and valid technique for speciation of a large number of samples in a timely manner. Previous work had shown the phenotypic hippurate hydrolysis test to be an unreliable indicator of species due to the presence of atypical hippurate negative *C. jejuni* strains (Totten *et al.*, 1987). This was corroborated by comparative studies in the CRU (Thwaites *et al.*, unpublished). Additionally the 16S rRNA genes and 23S rRNA genes had been extensively used for identification (Eyers *et al.*, 1993; Linton *et al.*, 1996) however data had suggested that there was insufficient sequence variation within these two genes to discriminate the two closely related species (Burnett *et al.*, 2002) and the contamination issues in using the Taqman with 16S rRNA genes (Corless *et al.*, 2000). Success had been achieved in using the Denis *et al* (1999) conventional multiplex PCR based on the *mapA* and *ceuE* genes. These two genes had also been reported as reliable indicators of the two species by On and Jordan (2003) and had also been used in the analysis of food samples (Denis *et al.*, 2001).

The *mapA* (membrane associated protein) was described by Stucki *et al* (1995). The gene was characterised through investigations of the unique differences in surface structure between *C. jejuni* and *C. coli*. The function could not be established, although the end of an Open Reading Frame upstream of the *mapA* gene shared significant homology with the 3' end of the *E.coli lepA* gene, (Salyers & Witt, 1994) which encoded for a membrane associated GTP binding protein. A high degree of

conservation of the *mapA* gene in *C. jejuni* and its apparent specificity for the species made it a good candidate for differentiation of *C. jejuni* and *C. coli.*

The *ceuE* (campylobacter enterochelin uptake) locus of campylobacter encodes for a polypeptide which functions as an iron chelating protein involved in siderophore transport (Section 1.7) The *ceuE* genes of *C. coli* and *C. jejuni* are, in some regions sufficiently divergent, making the *ceuE* gene of *C. coli* a suitable candidate gene for differentiation (Denis *et al.,* 1999).

These two genes had been described already by Denis *et al.* (1999) and the existing conventional PCR (and gel electrophoresis) was already in use in the CRU laboratory. However this was not suitable for high throughput screening so an assay to differentiate between the two species using the Taqman was devised. The development of a real time PCR assay to successfully speciate *C. jejuni* and *C. coli* based on the Taqman but also with the application to the Roche Lightcycler real time platform is described in this chapter. It was intended that this would provide a more rapid and robust method for the speciation of the two commonest campylobacters.

The assay has been applied to large sample numbers as encountered within the CRU laboratory and to other sample types including charcoal transport swabs. Additionally the assay has been extended to show detection from potential sources of campylobacter infection including poultry and meat samples. These applications involved the use of different DNA extraction methods, which during the course of this investigation were evaluated for use in later studies.

4.2 Materials and Methods

4.2.1 Design of Primers and Probes

Materials

- Primer Express Version 1.5 computer software (Applied Biosystems Warrington, UK.)
- Genbank *map*A sequence X80135
- Genbank *ceu*E sequence X88849

Method

The Primer Express program was used to design the primer and probe sequences. This program selects primers with optimised melting temperatures of 58-60°C, and probes within the range 68-70°C for successful amplification under universal cycling conditions. Other criteria were also met:

For the primers

- 20-80% GC content
- Length 9-40bp
- <2% difference in T_m between two primers
- Maximum of 2/5 G or C at 5' end

For the probes

- 20-80% GC content
- length 9-40bp
- no G on 5' end

- <4 contiguous genes
- Not more Gs than Cs

For the Amplicon

- 50-100bp in length
- As close as possible to the probe

(Applied Biosystems Taqman 7700 Training Materials)

The specificity of primer and probe sequences was confirmed by homology searches in BLASTn (www.ncbi.nlm.nih.gov/BLAST). Also, visual checks were made along the sequences to ensure that primer/dimer formation was unlikely between the two sets of primers and probe, within the duplex reaction. The Taqman probes were custom synthesized by Applied Biosystems and included two fluorescent dye extensions, the 3' quencher TAMRA (6-carboxy-tetramethyl-rhodamine) and a 5' reporter dye. Various different 5' reporter dyes were described in the literature from Applied Biosystems, however for a duplex reaction the reporter dyes FAM (6-carboxy-fluorescein) and VIC (Applied Biosystems license) were described as the most robust with a sufficient difference of 40nm between the two wavelengths, allowing for successful multiplexing.

4.2.2 Taqman PCR

Reagents for the Taqman PCR were supplied as a ready prepared mastermix where only the addition of primers and probes was required. The pre prepared mastermix was optimised with the correct buffer and hot start *Taq* polymerase, optimal $MgCl_2$ (15mM), the passive reference dye (ROX), and the inclusion of the enzyme

AmpErase UNG (Uracil N glycosylase) with dNTPs containing dUTP. This system was in place to prevent the re amplification of previous PCR products, the incorporation of dUTP into amplicons enabled the digestion by the enzyme AmpErase UNG.

Materials

- 1x Taqman Universal PCR Mastermix (containing AmpliTaq Gold™ DNA Polymerase, AmpErase UNG, dNTP's with dUTP, Passive reference 1 (ROX) and 15mM $MgCl_2$ (Applied Biosystems, Warrington. UK).
- Primers *map*A forward and reverse 3000nM stock (Table 4.4) (MWG Biotech)
- Primers *ceu*E forward and reverse 3000nM stock (Table 4.4) (MWG Biotech)
- Probe *map*A 1000nM stock (Table 4.4) (Applied Biosystems)
- Probe *ceu*E 1000nM stock (Table 4.4) (Applied Biosystems)
- Non template control (sterile distilled water) (Promega)
- Positive control (Equimolar concentrations of *C. jejuni* and *C. coli* DNA)
- 96 well Taqman PCR plate and adhesive Lid (Applied Biosystems)

Method

Each 25µl reaction mix contained, Taqman pre made mastermix, two sets of 300nM forward and reverse primers, two 100nM probes labelled with the dyes FAM (*mapA*) or VIC (*ceuE*), both with quencher TAMRA and 5µl of DNA. All reactions were carried out with a Non Template Control (NTC) and positive control in 96 well optical reaction plates and sealed with optical adhesive covers. The Taqman universal PCR cycling conditions were applied (Table 4.1).

Table 4.1 Taqman universal cycling conditions

	Temperature (°C)	Time (min:sec)	Activity
Step 1	60	2:00	Activation of UNG
Step 2	95	10:00	Activation of *Taq* DNA Polymerase
40 cycles of	95	00:15	Denaturaton
	60	1:00	Annealing/Extension

4.2.3 Data Analysis

Method

Reactions were programmed and data analysed on a Macintosh 4400/20 (Apple computer, Santa Clara, California) linked directly to the ABI Prism 7700 Sequence detection System using the SDS 1.7 application software (Applied Biosystems Manual). PCR products were detected directly by monitoring the increase in fluorescence. Real time analysis monitored the accumulation of the fluorescence throughout the 40 cycles producing an amplification plot over the entire course of the reaction (Figure 4.1). Results were displayed as a C_T (Threshold cycle) and ΔR_n.

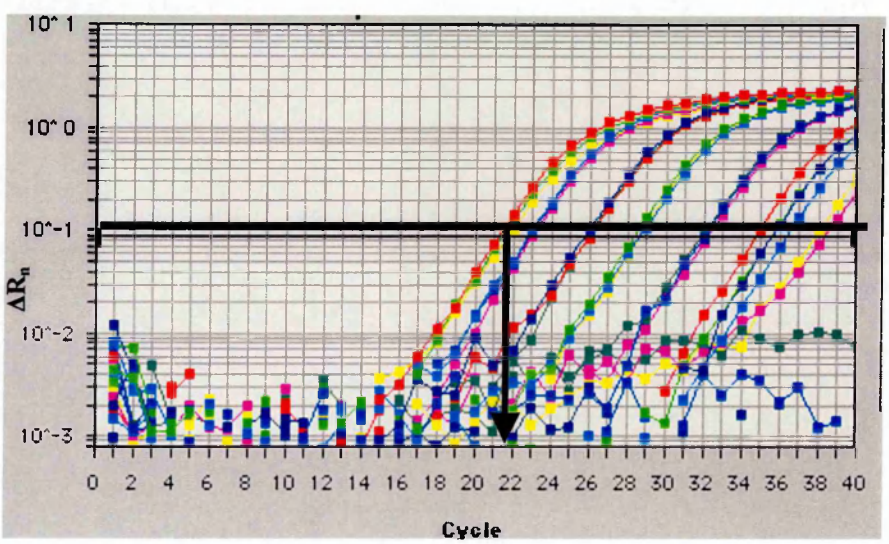
- C_T indicated the cycle number at which an increase in fluorescence was detected above a baseline level. For the NTC this was 40 (the total number of cycles) due to there being no template therefore no amplification occurring, hence a flat line. For positive samples the C_T was around 15-26 the cycle at which amplification and hence fluorescence started to accumulate and be detected (Taqman training materials, Applied Biosystems).

- ΔR_n referred to the normalised reporter signal R_n (fluorescence signal of the reporter dye divided by the fluorescence of the passive reference dye (ROX))

minus the baseline signal established in the few first cycles of the PCR, this increased during the reaction until it reached a plateau (Taqman training materials, Applied Biosystems).

The threshold was set manually so that it intersected half way across the amplification plots. Results in the form of C_T values were exported into an Excel spreadsheet.

Figure 4.1 Example of a typical Taqman amplification plot showing cycle number plotted against ΔR_n. The horizontal line across the plot was a set threshold, which was adjusted manually to intercept the mid point of the amplification plots. The threshold cycle (C_T) indicated by the arrow shows the cycle number at which each plot crossed the set threshold line.



4.2.4 Optimisation of primer and probe concentration

Different concentrations of forward and reverse primers for each gene sequence were evaluated separately using varying final concentrations of 50nM, 300nM and 900nM forward or reverse primer per PCR reaction with standard cycling conditions, standardised concentrations of probe (100nM) and the recommended amount of *C. jejuni* or *C. coli* DNA (10ng per reaction) by Applied Biosystems. Primer optimal reaction concentrations were determined in each case from the least amounts of primer added producing the lowest C_T value and the highest ΔR_n. The effect of varying the concentration of probes was also determined from 50nM to 250nM (in 50nM intervals) in a similar manner and the effect on the C_T and ΔR_n was determined.

Coamplification of the two targets was also investigated. Spiking experiments using triplicate concentrations of equivalent DNA from both species were tested with the duplex assay. Due to the quantitative nature of the Taqman assay, if one reaction had a significantly lower C_T than the other it could be assumed that the reactions were not operating equally, which would require the addition of more (or less) of one of the primer and probe sets. Usually duplex reactions require primer limiting experiments to be applied due to the requirement for both gene sequences to be positive, however this reaction functioned with only one gene sequence required to be positive at one time, the chance of both *mapA* and *ceuE* being positive was unlikely.

4.2.5 Reduction in total reaction volume

The recommended total reaction volume as suggested by Applied Biosystems was 50µl. The effect of reducing the reaction volume by half to 25µl was investigated, primarily to save on reagent costs. The assay was applied using three isolates of each species and a NTC. One set of reactions was made up in full 50µl reaction volume, and a second set of duplicates was scaled down to a volume of 25µl.

4.2.6 Sensitivity of the primers and probes

Sensitivities were determined separately for amplification of each gene region by the method described by Applied Biosystems (Anon, 1997). Standard concentrations of *C. jejuni* and *C. coli* purified DNA were prepared by a MagNApure extraction (Section 2.4.4). These were RNAse treated (Section 2.4.6) and the concentration of DNA determined by use of a Biophotometer (Eppendorf). Each suspension was divided into three aliquots and 10-fold serially diluted to produce a range of dilutions from 10ng/µl to $1x10^{-8}$ ng/µl. These were used in the Taqman assay and a standard curve was plotted of C_T against amount of DNA to determine the mean limits of detection for each gene.

4.2.7 Specificity of the Assay

Materials

- NCTC *Campylobacter* species (Table 4.4) (National Collection of Type Cultures, London UK.)

Method

DNA was extracted from fifteen National Collection of Type Cultures (NCTC) strains of *Campylobacter,* and nine other related organisms and tested with the Taqman assay.

4.2.8 Comparison of DNA extraction methods from cultured cells

Materials

- DNA extracted by 3 methods (Sections 2.4.1 and 2.4.4)
  - (i) MagNApure Total NA Isolation Kit III
  - (ii) MagNApure Bacterial DNA Isolation Kit
  - (iii) Boiled Cell Lysates

Method

A 24-hour culture of *C. jejuni* was prepared to 0.3 at OD_{600} in SDW and divided into three 500µl aliquots. These were 10 fold serially diluted in SDW to a total volume of 100µl. Different DNA Extraction methods were compared on the set of serially diluted cells, and the lowest culture dilutions, which could be detected by the Taqman PCR assay, were determined. The method using cell lysates was used in all subsequent experiments due to ease of use and adequate DNA recovery for use in this assay.

4.2.9 Large scale use of the assay

This purpose of this section was to assess the viability of using the assay in a routine work setting for the speciation of *C. jejuni* and *C. coli* strains received into the CRU

over a six-month period. (This section was carried out with the help of the staff of the CRU; I especially wish to acknowledge the help of Ella Powell in this section.)

Materials

- Strip of 8 cupules and lids (AB gene)
- Reagents as before (Section 4.2.2)

Method

The technique was applied as described in section 4.2.2 except for the arrangement of cell lysates within strips of 8 cupules with caps to allow for easy transfer of samples using multi-channel pipettes.

The assay was utilised over a six-month period, during which all 6015 isolates speciated were also sero and phage typed. Any isolates failing on the Taqman assay were re-tested and subsequently tested by a conventional PCR for other possible species of campylobacter, or subjected to a conventional PCR for *C. jejuni/C. coli* differentiation as described by Vandamme *et al* (1997).

4.2.10 Adaptation of the assay to the Lightcycler system

The Lightcycler (LC) is an alternative platform for real time rapid PCR which tends to be more widely available in laboratories due to the lower cost and versatility of the machine. The purpose of this section was to investigate the application of the Taqman assay described, onto the LC platform for use in laboratories where the Taqman system was unavailable.

Two formats were tested on the Lightcycler:

- Use of the Taqman primers (no probes) with the sequence independent binding dye SYBR Green
- Use of Taqman primers and probes

Taqman hydrolysis probes can be used on the LC, however the LC has set channels for detection of specific wavelengths, designed for the use of LC hybridisation probes or fluorescein these being at 540nM, 640nM and 705nM. The Taqman hydrolysis probes described here, had emission wavelengths of 520nM (FAM) and 560nM (VIC) therefore could only be detected within the 540nM channel, meaning that duplexing the two probes was impossible.

4.2.11 Lightcycler PCR using SYBR Green.

Materials

- Faststart DNA Master SYBR Green I Kit (Roche Diagnostics Ltd, Lewes, UK)
- Taqman primers (Table 4.4) (MWG Biotech)
- 15mM $MgCl_2$ (Roche Diagnostics)
- Sterile distilled water (Roche Diagnostics)
- Glass reaction capillaries and plastic stoppers (Roche Diagnostics)

Method

A duplex reaction could not be applied due to the use of SYBR green, therefore for every sample two reactions had to be set up (1 for *mapA* and 1 for *ceuE*). The LC Faststart DNA Master SYBR Green I Kit was used with the Taqman primers. Each 20µl reaction contained 11.6µl PCR grade distilled water, 4mM $MgCl_2$, 0.5µM each primer, 2µl LC Faststart DNA master SYBR Green 1, and 2µl DNA. Reactions were prepared in LC capillaries, and sealed with the plastic stoppers. The capillaries were centrifuged at 3000rpm for 10 seconds to bring down contents and placed into the carousel of the Lightcycler 1.2. Cycling conditions were used as shown (Table 4.2), all data was analysed by the Lightcyler instrument with version 3.5 software.

Table 4.2 Lightcycler cycling conditions

Program	Temp (°C)	Time (min:sec)	Temp transition rate (°C/sec)	Analysis mode
Denaturation	**95**	**3:00**	**20:00**	**None**
Amplification (45 cycles)	**95**	**15**	**20:00**	**None**
	59	**5**	**20:00**	**none**
	72	**10**	**20:00**	**single**
Melting Curves	**95**	**0**	**10:00**	**None**
	65	**15**	**20:00**	**None**
	95	**0**	**0.1**	**Cont**
Cooling	**40**	**30**	**20:00**	**None**

The fluorescence parameters were set as described in the manufacturers instructions and the F1 channel was used to monitor fluorescence.

4.2.12 Lightcycler PCR using Taqman probes.

Materials

- Faststart DNA Master SYBR Green I Kit (Roche Diagnostics Ltd, Lewes, UK)
- Taqman primers (Table 4.4) (MWG Biotech)
- 15mM $MgCl_2$ (Roche Diagnostics)
- Sterile distilled water (Roche Diagnostics)
- Glass reaction capillaries and plastic stoppers (Roche Diagnostics)
- Taqman Probes (Table 4.4) (Applied Biosystems)

Method

Two single reactions were prepared, 1 each for *mapA* and *ceuE,* containing 2µl LC faststart hybridisation probes mix, 0.4µM probe, 0.1µM each primer and 4mM M_gCl_2. Reactions were performed in glass capillaries as described above with the following cycling conditions (Table 4.3).

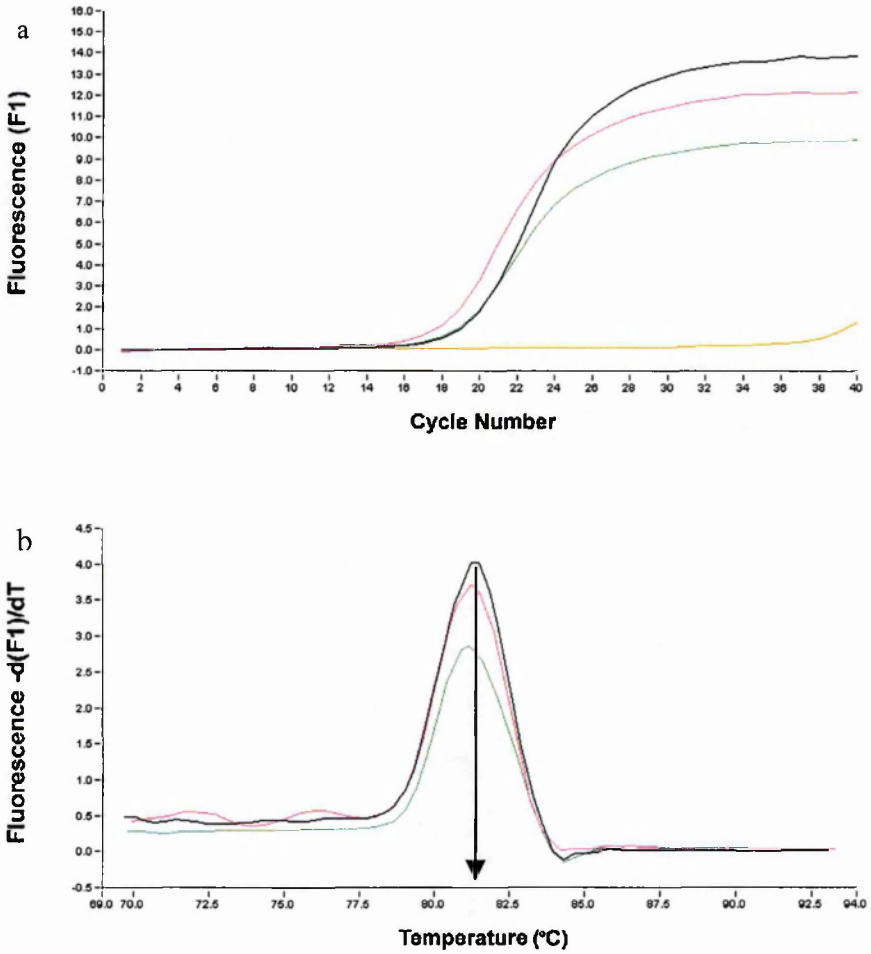
Table 4.3 Lightcycler cycling conditions forTaqman probes

Program	Temp (°C)	Time	Temp transition rate	Analysis mode
Denaturation	95	3:00	20:00	None
Amplification (40 cycles)	95	15	20:00	None
	60	5	20:00	none
	72	10	20:00	single
Cooling	40	30	20:00	None

4.2.13 Optimisation of Lightcycler PCRs

In each LC format, reactions were optimised with respect to the annealing temperature, concentrations of primers and probes and $MgCl_2$ was titrated from 1-5mM per reaction. Results were read from the amplification plot where a positive sample had a crossing point of 10-25 (Figure 4.2a). Negative samples were usually in the range of 35-40. All amplifications with SYBR green were validated by the melt curve data, which was performed after the amplification stage. This step slowly heated up the DNA samples in order to melt the DNA. The temperature at which the DNA melted was characteristic to its sequence depending on the length and the order of bases, hence all the *mapA* and all the *ceuE* amplified products melted at characteristic temperatures (81.5°C) (Figure 4.2b). The LC reactions using the probes could not be analysed with melt curve data, due to the release of the fluorophore freely into solution, therefore melting amplified products would not result in a decrease in fluorescence emissions.

Figure 4.2 Example of LC Amplification with SYBR green (a) and melt curve data (b). Graph A shows 3 typical amplification plots and one negative sample (yellow line). Graph B shows typical melt curve data, three samples showing similar sized peaks with temperatures of 81.5°C indicative of similar PCR product amplification.



4.2.14 Application of the Assay directly to Charcoal Transport Swabs

Materials

- MagNApure Total Nucleic Acid Kit III
- Distilled water

Methods

A total of 30 randomly selected isolates received into the laboratory on swabs were subjected to DNA isolation directly from the swab material using the MagNApure (Section 2.3.2). In parallel, standard inoculation of the swab onto CBA plates was carried out followed by 24 hours incubation and the preparation of a boiled cell lysate. Both sets of 30 isolates were analysed with the Taqman assay.

4.2.15 Comparison of DNA Extraction Methods for Direct Use of the Assay on food samples.

Materials

DNA was isolated by three methods from artificially contaminated meat and milk samples (Section 2.3.3). Only the MagNApure Bacterial DNA Isolation Kit was tested in this section due to its superior results over the MagNAPure Total Nucleic Acid Isolation Kit III (as described in Section 4.3.6).

(i) MagNApure Bacterial DNA Isolation Kit

(ii) "Bugs 'n Beads" (Genpoint)

(iii) Prepman Ultra™ Reagent

Food samples

- Poultry (including fresh chicken, turkey and duck),
- Fresh pork, lamb and beef
- Fresh pasteurised semi skimmed milk

Method

Food samples (meat rinses and milk) were artificially inoculated with 24-hour cultures of *C. jejuni* or *C. coli* as (Section 2.3.3). DNA was isolated from the dilutions of cells in meat rinses or milk using three methods, the MagNApure (Section 2.4.4), the Bugs 'n Beads kit and the Prepman Ultra Reagent (Applied Biosystems) (Sections 2.5.1 and 2.5.2).

4.2.16 Direct application of the assay to naturally contaminated food samples

Methods

All meat samples (5g) were added into 15mls campylobacter enrichment broth, shaken vigorously for 2 minutes and vortexed for 2 minutes to release the cells into solution. The contents were allowed to settle and a 200µl sample of the rinse fluid was removed. Two replicates of 100µl were used in a direct MagNApure extraction (Section 2.4.4) using the Bacterial DNA Isolation Kit and then tested with the Taqman assay. The remainder of the rinse fluid was incubated for enrichment and cultured (Section 2.3.3) to provide confirmatory culture results.

4.3 Results

4.3.1 Design of primers and probes for the Taqman speciation assay

The published sequences were obtained from GenBank (GenBank accession numbers, X80135 (*mapA*) and X88849 (*ceuE*)). The sequence for *mapA* was imported into Primer Express and a region between 305 and 400bp was identified where the optimal design parameters for primers and probe could be met. A *ceuE* gene was present in both *C. jejuni* and *C. coli* therefore sequences from each (X88849 and X82427) were aligned using the clustalW multiple alignment algorithm in Bioedit and species-specific areas were identified. These were used to discriminate between the two species. A region between 745 and 850 base pairs (bp) was identified using Primer Express, which encompassed divergence between the two species whilst adhering to the conditions specified by Primer Express (Appendix 1.1) (Table 4.4). The homology search in BLASTn revealed no other identical sequences than those reported for the *map*A or *ceu*E genes.

Table 4.4 Primer and probe sequences for *mapA* and *ceuE* gene regions

Gene	Primer/probe	Sequence (5'-3')	Amplicon size (bp)
mapA	*mapA* forward	CTGGTGGTTTTGAAGCAAAGATT	95
	mapA reverse	CAATACCAGTGTCTAAAGTGCGTTTAT	
	mapA FAM probe	TTGAATTCCAACATCGCTAATGTATAAAAGC CCTTT	
ceuE	*ceuE* Forward	AAGCTCTTATTGTTCTAACCAATTCTAACA	102
	ceuE Reverse	TCATCCACAGCATTGATTCCTAA	
	ceuE VIC probe	TTGGACCTCAATCTCGCTTTGGAATCATT	

4.3.2 Optimisation of the reaction

(i) Primer and probe concentrations

Optimal primer concentrations were determined from the mean results (n=5) showing the lowest C_T and highest ΔR_n. For the *mapA* reaction the forward and reverse primers amplified optimally at an equimolar concentration of 300nM with a mean C_T of 18.6 and a ΔR_n of 2.2 (Table 4.5a). For the *ceuE* reaction, optimal forward and reverse primer concentrations were also determined at an equimolar concentration of 300nM giving a C_T of 21.2 and ΔR_n of 1.4 (Table 4.5b). Optimal probe concentrations for each different reaction were determined to be 100nM (Table 4.5c). For *mapA* this concentration of probe gave a C_T of 25.6 and ΔR_n of 0.9, and for *ceuE* a C_T of 26.3 and ΔR_n of 1.1 (Table 4.5c). In the presence of lower concentrations of probes false negative results were shown due the limiting effect of insufficient reaction components. (Final optimal reaction concentrations are shown in Table 4.5d)

Table 4.5a Mean optimal Primer concentrations determined experimentally for the *mapA* gene region. Mean C_Ts determined with variation in each primer concentration. Highlighted box showing optimal concentration (black –C_T, red-ΔR_n).

Primer concentration (nM)		Forward primer concentration		
		900	300	50
Reverse primer concentration	900	18.7 2.4	22.9 1.32	23.9 1.19
	300	27.1 1.29	18.6 2.2	23 1.69
	50	22.4 1.31	31.6 1.24	19.2 0.94

Table 4.5b Mean optimal Primer concentrations determined experimentally for the *ceuE* gene region. Mean C_Ts determined with variation in each primer concentration. Highlighted box showing optimal concentration (black –C_T, red-ΔR_n)

Primer concetration (nM)		Forward primer concentration		
		900	300	50
Reverse primer concentration	900	26.3 1.1	27.1 0.8	27.2 0.8
	300	27 0.8	21.2 1.4	27 0.8
	50	26.8 0.9	27.2 0.8	29 0.3

Table 4.5c Mean optimal probe concentration determined experimentally for the *mapA* gene region. Mean C_Ts calculated with variation in probe concentration. Highlighted box showing optimal concentration

Probe Concentration (nM)	*map*A		*ceu*E	
	C_T	ΔR_n	C_T	ΔR_n
250	20.6	0.8	16	0.6
200	21.4	0.9	25	1.4
150	25.9	0.56	26	1.2
100	25.6	0.9	26.3	1.1
50	36.3	0.03	28	0.37

Table 4.5d Optimal primer and probe concentrations for each gene region

Gene sequence	Forward primer (nM)	Reverse primer (nM)	Probe (nM)
mapA	300	300	100
ceuE	300	300	100

(ii) Optimisation of the duplex reaction

Spiking experiments demonstrated both gene regions could be co-amplified simultaneously and equal C_Ts were determined in each case (Table 4.6). Therefore this assay was able to detect mixed infections. Very little re-optimisation was actually required to facilitate co-amplification of the two gene regions, in each case the primer and probe concentrations remained at the optimal levels.

Table 4.6 Comparison of C_T values for co-amplification of each gene region, with spiked DNA samples

DNA concentration (ng/µl)		C_T value *mapA* gene	C_T value *ceuE* gene
C. jejuni (11168)	*C. coli* (13050)		
20	20	15.6	16.1
10	10	17.6	17.8
5	5	19.4	20.1

(iii) Reduction in Total Reaction Volume

Reaction efficiency was not compromised with the reduced reaction volume and there was very little change seen in the value of each of the C_Ts when identical isolates were compared (Table 4.7).

Figure 4.7 Effect on C_T by the scaling down of reaction volume from the recommended 50µl to 25µl.

Isolate (species)		*mapA* gene		*ceuE* gene	
		50µl Reaction volume (C_T)	25µl Reaction volume (C_T)	50µl reaction volume	25µl reaction volume
(*C. jejuni*)	207	20.4	20.6		
	11168	19.8	19.9		
		21.7	21.9		
(*C. coli*)	206			21.6	21.7
	20969			19.2	19.3
				15.2	15.3
NTC		40	40	40	40

4.3.3 Sensitivity of the Assay

Standard curves were plotted using the mean of three replicate dilution series for each gene region. C_T value against log DNA concentration (pg) was plotted where a linear relationship was obtained (Figures 4.3 and 4.4). Detection limits in genome copies were determined using the conversion factor of the weight of 1 genome of *C. jejuni* being equal to 1.6fg DNA (as determined by Sails 2002). This value was also used for the *ceu*E gene region due to the absence of an accurate value for the size of the *C. coli* genome.

Calculations to determine the input amount of DNA were as follows (Anon, 1997). The slope of the line was determined in Microsoft Excel in the format $y = mx + b$. Where b indicated the point at which the standard curve crossed the y axis (red line). The value m indicated the gradient of the standard curve as calculated in Excel. To determine the detection limits of the assay (minimum input of DNA) a C_T value of 39 was taken as being positive and 40 was considered negative. By determining the value of x in the equation it was possible to determine the amount of DNA in the reaction, when a C_T value of 39 was obtained. From this value it was possible to determine the number of genome copies by using the value of 1.6fg, the weight of one *C. jejuni* genome (as calculated by Sails 2002).

For the *map*A region
Line $y = mx + b$
Line $y = -3.5224x + 33.586$
([cell containing C_T value]-b)/m
39-33.586/-3.5 = -1.54
10^ [cell containing log input amount]
10^ (–1.5)
= 0.02 pg (20fg)
0.02/1.6 (weight of one *C. jejuni* genome (fg))
12.5 genome copies
For the *ceuE* region
Line $y = -3.474x + 37.08$
39-37.08/-3.47 = -0.55
10^ (–0.55)
= 0.27 pg (279fg) /1.6 (weight of one genome)
174 genome copies

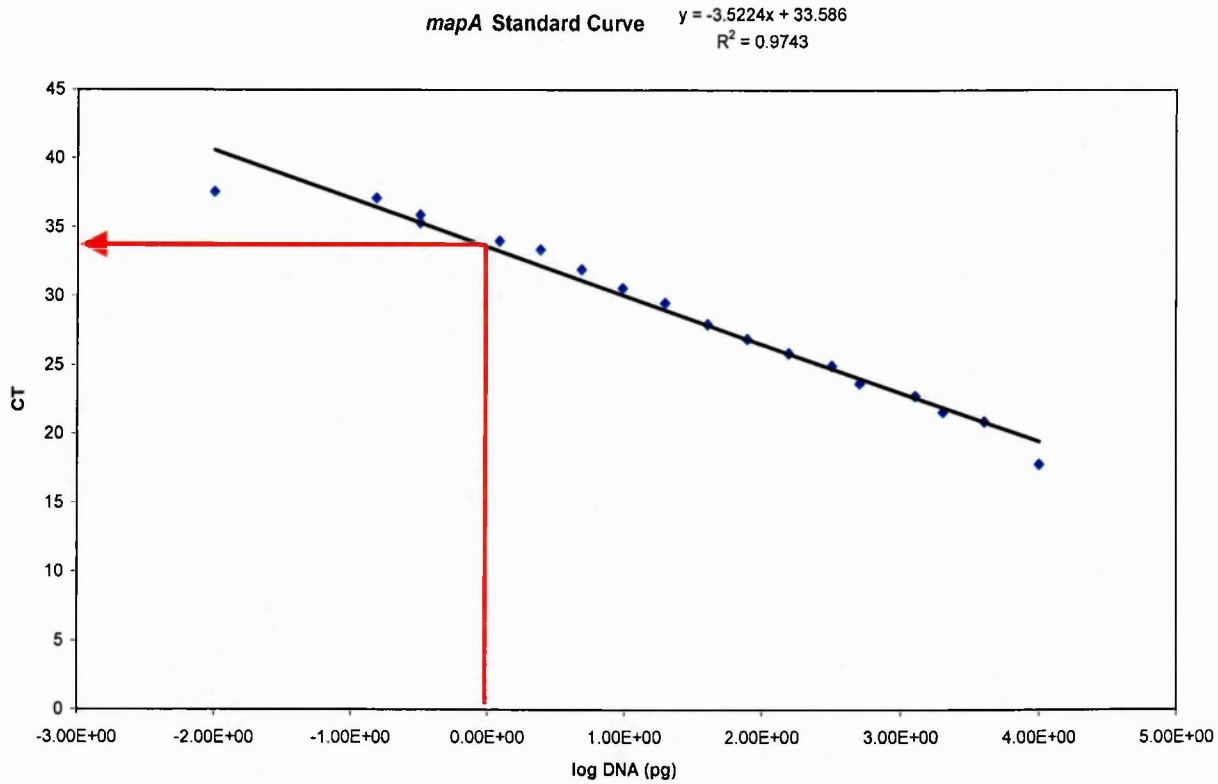


Figure 4.3 Standard curve for the *map*A target detected using a FAM labelled probe. Calculations enabled a detection limit when using a C_T value of 39 of 20fg of DNA or 12.5 *C. jejuni* genome copies.

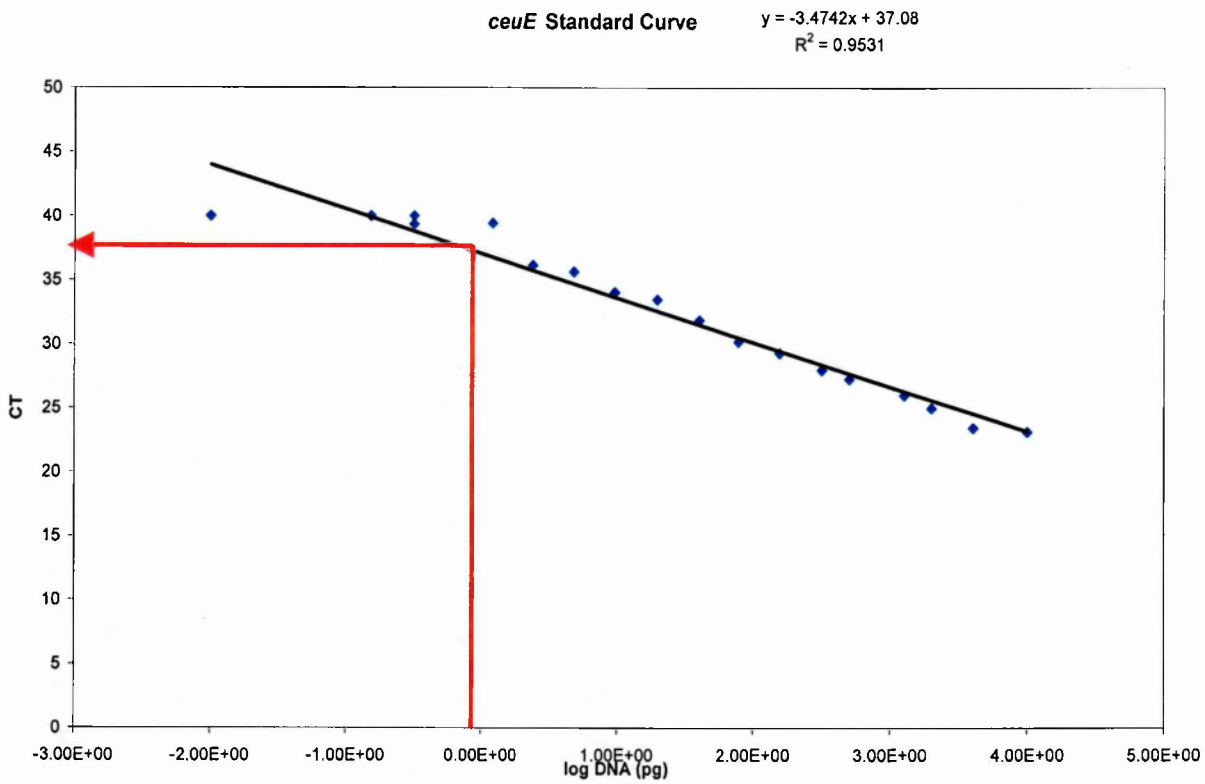


Figure 4.4 Standard curve for the *ceu*E target using a VIC labelled probe. Calculations enabled a detection limit when using a C_T value of 39 of 279fg of DNA or 174 *C. coli* genome copies.

4.3.4 Specificity of the assay

For every organism tested other than *C. jejuni, C. jejuni* subsp. *doylei* and *C. coli* results were negative and no false positives were detected. It was particularly important for the other species of *Campylobacter* that no false positives occurred as these would be the most likely organisms to be isolated in the place of *C. jejuni* or *C. coli.* Similarly *Arcobacter* and some species of *Helicobacter* may be potentially isolated due to having similar growth requirements (Table 4.8).

Table 4.8 Specificity of the assay using DNA from other microrganisms and *Campylobacter* species (shaded areas indicating positives)

Genus	Species	NCTC/ref no.	Source	Taqman Result (C_T)	
				mapA	*ceuE*
***Campylobacter* species**	*C. jejuni ss jejuni*	NC11351-01	NCTC	17.4	40
	C. jejuni ss *doylei*	NC11847-04	NCTC	20.7	40
	C. coli	NC11353-04	NCTC	40	22.1
	C. fecalis	NC11415-04	NCTC	40	40
	C. lari	NC11352-07	NCTC	40	40
	C. consiscus	NC11485-06	NCTC	40	40
	C. fetus	NC10842-07	NCTC	40	40
	C. fetus	NC10354-07	NCTC	40	40
	C. hyointestinalis	NC11608-06	NCTC	40	40
	C. helveticus	NC12470-03	NCTC	40	40
	C. curvus	NC11649-03	NCTC	40	40
	C. sputorum ss *mucosalis*	NC11000-05	NCTC	40	40
	C. sputorum ss bubulus	NC11367-03	NCTC	40	40
	C. sputorum	NC11528-05	NCTC	40	40
	C. upsaliensis	NC11541-05	NCTC	40	40
Other organisms	*E. coli*	NC09001-19	NCTC	40	40
	S. typhi	ATCC-4543	NCTC	40	40
	A. butzleri	*NC12481-03*	*NCTC*	40	40
	A. cryaerophilus	*NC11885-04*	*NCTC*	40	40
	A. skirrowski	*NC12713-02*	*NCTC*	40	40
	A. nitrofigilis	*NC12251-02*	*NCTC*	40	40
	H. pylori	NC11637	LEP	40	40
	H. pullorum	C77339	LEP	40	40
	H. canadensis	C78661	LEP	40	40

*suffix indicates batch number

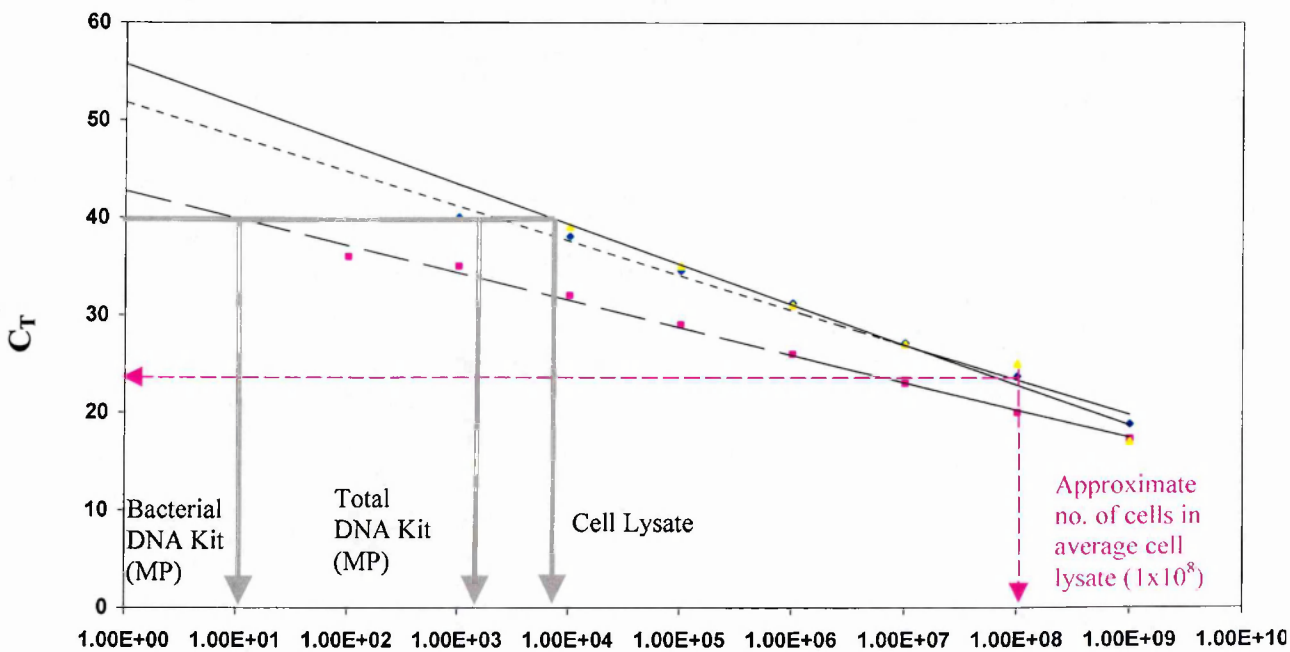
4.3.5 Comparison of different DNA extraction methods

Three different DNA extraction methods were compared by the use of triplicate 10-fold serial diluted cultures. Comparisons using the number of colony forming units were used in this instance due to difficulties in quantifying the amount of DNA in a cell lysate preparation. The method utilising the MagNApure in conjunction with the Bacterial DNA isolation kit resulted in the most DNA isolated with the lowest mean C_T for the initial dilution of $1x10^9$ cells (C_T of 16.2). The recovery rate with this method for the Taqman assay was $1x10^1$ CFU/ml (Table 4.9). Even though the recovery rate for the cell lysates was lower at $1x10^4$ CFU/ml, this method was used in all subsequent experiments due to the ease and speed of preparation (approx. 20 minutes) and lower cost (approx £0.05 per sample). The mean number of CFU/ml was determined for preparation of cell lysates (1 colony within 100µl SDW) by construction of a standard curve (Figure 4.5). Thus it was possible to determine that a mean C_T of 24 from a standard cell lysate preparation was equivalent to $1x10^8$ CFU/ml, sufficiently above the recovery rate, therefore suitable for further experiments.

Table 4.9 Comparison of different DNA extraction methods by CFU/ml, cost and time *(2002 prices)

Method of Extraction	Approximate time taken (per 32 samples)	Approximate cost per sample *	Mean minimum levels of *C. jejuni* cells determined CFU/ml
MagNAPure Bacterial DNA Isolation Kit	40 minutes preparation 90 minutes machine running time Total 2 hours 10 minutes	£1.70	1×10^1
MagNApure Total NA Isolation Kit III	20 minutes preparation 90 minutes machine running time Total 1 hour 50 minutes	£1.70	1×10^3
Cell Lysates	10 minutes preparation 10 minutes boiling Total 20 minutes	£0.05	1×10^4

Figure 4.5 Comparison of Extraction Methods with estimated detection limits for the three different extraction methods (MagNApure (MP) Bacterial Kit (---), MagNApure Total Kit (''''), Cell Lysates(—)) from cultured cells. Dashed line (---) showing approximate number of cells within a cell lysate (Mean C_T of 5 cell lysates and estimated from the standard curve).



4.3.6 Application of the assay to large sample numbers

During six months of using the Taqman assay, 5877 out of 6015 (97.6%) cell lysates tested were either *mapA* positive or *ceuE* positive, therefore identified as being either *C. jejuni* or *C. coli* (Table 4.10). Of these 5336 (88.7%) were identified as *C. jejuni* (*mapA* positive/*ceuE* negative) and 541 (9%) identified as *C. coli* (*mapA* negative/*ceuE* positive). In all strains the Taqman molecular identification was consistent with the species specificity of the serotype and any non-serotypeable strains were confirmed using the conventional PCR. Of the remainder, 123 (2.1%) failed to amplify with either gene. Ninety-five (1.6% of total) of these samples were identified by the Vandamme PCR to be *C. jejuni*, 17 (0.3% of total) were identified by the same technique to be *C. coli.* Both of the above were also confirmed by serotype. Thus 1.9% of isolates were only identified by the conventional PCR of Vandamme *et al* to be *C. jejuni/coli.*

Table 4.10 Species determination of 6015 strains.
[a] **Confirmed by serotype (Frost *et al.*, 1998)**
[b] **Identified by conventional PCR (Vandamme *et al.*, 1997)**
[c] **Identified by 16S rRNA RFLP(Cardarelli-Leite *et al.*, 1996)**

Gene Sequence Amplified	No. (%)	Species Identification
***mapA* +/*ceuE* -**	5336(88.7%)	*C. jejuni*[a]
***mapA*-/*ceuE*+**	541(9%)	*C. coli*[a]
***mapA*-/*ceuE*-**	123(2.1%)	95 (1.6%) *C. jejuni*[b] 17 (0.3%) *C. coli*[b] Other species 11 (0.2%)[c] 2 *C. upsaliensis* 1 *C. fetus* 8 *C. lari*
***mapA*+/*ceuE*+**	15(0.2%)	10 (0.16%) true mixed colonies 5 (0.08%) potential "hybrids".
Total	6015	

4.3.7 Investigation of potential "mixed" strains

A small number of isolates (n=15) possessed both genes, and further tests were carried out to try to resolve them into a single species. Initially this involved multiple inoculations of distinct single colonies onto separate media plates, to eliminate the possibility of genuine mixes on culture plates. Ten of the fifteen mixtures gave separate colonies belonging to one or other of the species following these single colony picks. Five potential "mixed isolates" remained. For these remaining isolates, further investigations were carried out. Conventional PCR using the primers as described by Denis (Denis *et al.,* 1999) for the *mapA* and *ceuE* genes were applied separately and in each case products of the correct size were obtained. These were sequenced in both directions, using the same primers (Section 2.3). Contigs were assembled and edited in Bioedit and aligned using the clustalW multiple alignment function against the relevant Genbank published sequences for the *mapA* (X80135) or *ceuE* genes (X88849) as well as control *map*A and *ceu*E positive isolates 206 (*C. coli*) and 207 (*C. jejuni*) (Appendix 1.2 and 1.3) The *map*A alignment showed very little divergence in the potential mixed strains when compared to the published sequence and the control strain 207. The small numbers of mismatches were in areas distinct from the primer and probe binding regions, the only region where a mismatch occurs in all except the published strain was at 278bp. Similarly, for the *ceu*E alignment no divergence was seen in the area of the probes or primers and generally throughout the region sequenced there was very little sequence variability except for a base change at 38bp. The sequencing confirmed the presence of both the gene regions within the potentially "mixed" strains. Investigations of the strain type confirmed that they were not from an unusual source or

4.3.8 Adaptation of the assay to the Roche Lightcycler real time PCR system

(i) SYBR Green sequence independent amplification

The $MgCl_2$ titration determined the optimal reaction concentration for both the *mapA* and *ceuE* primers to be 3mM. At this concentration optimal amplification occurred, with a mean crossing point of 17 for *mapA* and 18 for *ceuE*. Melting curve analysis determined the amplified products have melting temperatures of 81.5°C for *mapA* and 80°C for *ceuE* products.

When applied to 30 isolates results showed characteristic amplification plots for the *mapA* or *ceuE* amplified gene sequences with melting peaks of 81.5°C for *mapA* and 80°C for *ceuE*. All results were 100% concordant with the Taqman result.

(ii) Sequence Dependent Amplification with Taqman Probes

The optimal concentration of $MgCl_2$ was found to be 4mM for both *ceuE* and *mapA* reactions. At this concentration amplification plots showed a mean crossing point of 17.2 for *mapA* and 18.9 for *ceuE*. Results were 100% consistent with the results for the other formats (Table 4.11).

Table 4.11 Comparison of results from Taqman and the two different formats on the Lightcycler.

Gene Sequence Amplified	No of isolates positive by each method		
	Taqman	LC SYBR green	LC probes
mapA +/ceuE -	22	22	22
mapA-/ceuE+	8	8	8
Total	**30**	**30**	**30**

4.3.9 Application of the Assay to Charcoal Transport Swabs

Successful speciation was achieved directly from the charcoal transport swab with the use of the MagNApure for direct extraction. Results were concordant for both species. To determine the recovery rate would have been difficult, due to the non-standardised quantity of bacterial cells added initially, the semi solid nature of the transport media made it impossible to quantify (Table 4.12).

Table 4.12 Summary of results for 30 isolates with species determination by the Taqman assay based on standard boiled cell lysates and direct DNA isolation using the MagNApure

Species determination by Taqman	
Culture and boiled cell lysate preparation (no. of isolates)	**Direct DNA extraction from transport swab material (no of isolates)**
C. jejuni (27)	*C. jejuni* (27)
C. coli (3)	*C. coli* (3)

4.3.10 Application of the assay to potential sources of infection-artificially contaminated food samples

Three different genomic DNA extraction methods were compared for the direct isolation of campylobacter DNA from meat rinse samples and milk, which had been artificially contaminated with known concentrations of *C. jejuni* cells. Recovery of *C. jejuni* cells was possible with all three methods tested, however there were variations in rates of recovery. For each sample type tested the MagNApure DNA isolation kit resulted in the most DNA isolated with the lowest C_{Ts} in each case for the initial dilution of $1x10^8$ cells spiked into the meat rinse samples and milk. The recovery rate with this method for the Taqman assay was 100 CFU/ml (Table 4.13). The highest recovery rate was seen from the beef sample (100 CFU/ml) with the duck sample having the lowest recovery rate (1000 CFU/ml) with the MagNApure method. Lower recovery rates were seen with both other methods the Bugs 'n Beads and Prepman Ultra reagent where recovery was between 1000-10000 CFU/ml depending on the sample type. Despite the MagNApure having the best recovery rate, it was the slowest method, with the Prepman method taking 20-30 minutes and the Bugs 'n Beads 60-90 minutes. However due to the superior recovery rate from the food samples tested the MagNApure was used for further analyses.

Table 4.13 Comparison of different DNA extraction methods from food samples by CFU/ml and time. (Highlighted region showing the MagNApure with the Bacterial DNA Isolation Kit, which enabled the detection of lower levels from all sample types)

Extraction Method	Approximate time taken (per 32 samples)	Approximate cost (per sample)	Mean Minimum levels of *C. jejuni* cells determined from food samples CFU/ml						
			Meat Rinse samples						Pasteurised Semi skimmed milk
			chicken	duck	turkey	lamb	pork	beef	
MagNAPure Bacterial DNA Isolation Kit	20 minutes preparation 90 minutes machine running time **Total 2 hours 10 minutes**	£1.70	$1x10^2$	$1x10^3$	$1x10^2$	$1x10^2$	$1x10^2$	$1x10^1$	$1x10^2$
Bugs 'n Beads	**60-90 minutes**	£1.75	$1x10^4$	$1x10^4$	$1x10^4$	$1x10^4$	$1x10^4$	$1x10^3$	$1x10^3$
Prepman Ultra™ Reagent	**20-30 minutes**	£0.70	$1x10^4$	$1x10^4$	$1x10^4$	$1x10^4$	$1x10^3$	$1x10^3$	$1x10^4$

4.3.11 Application of the Assay to Naturally Contaminated Food Samples

Results from testing the naturally contaminated samples showed the presence of either *C. jejuni* or *C. coli* within 48% (n=35) of retail meat samples tested (Table 4.14). Of these positive results the majority were *C. jejuni* (43%) with only small numbers of *C. coli* isolated (5%). When the same samples were cultured following enrichment in Bolton broth only 28% were confirmed by growth on media. As expected the C_T results obtained for the Taqman assay were much higher than those obtained with extracted genomic DNA. The majority of samples, which were positive by Taqman and culture confirmed, were from chicken samples. No *C. jejuni* positive results (Taqman with culture confirmation) were obtained from any other type of meat sample including other poultry (duck), beef or lamb. The only *C. coli* positive result by Taqman and culture was obtained from a pork sample (Table 4.14).

Table 4.14 Results for naturally contaminated meat samples (n=35) tested by the Taqman assay and confirmatory result by culture.

Source	Sample type	Species by Taqman Assay (C_T value)	Culture confirmed
Poultry-Chicken samples	Breast-1	Cj (30)	Yes
	Legs-1	Cj (31)	Yes
	Legs (Halal)	Cj (35)	Yes
	Diced breast (low fat)	Cj (35)	Yes
	Breast (Corn Fed)	Negative	No
	Breast (Free range)	Cj (35)	No
	Thighs (boneless)	Cj (32)	No
	Thighs (on the bone)	Cj (35.5)	No
	Livers	Cj (33)	No
	Wings	Negative	No
	Legs (organic)	Cj (35)	Yes
	Breasts (Grade A)	Negative	No
	Breast (frozen)	Negative	No
	Legs (frozen)	Cj (32)	No
	Diced breast	Cj (31)	Yes
	Goujons	Cj (34)	Yes
	Fillets	Cj (38)	No
	Chicken breast skin	Cj (36)	Yes
Poultry-Turkey samples	Steaks	Cj (31)	Yes
	Mince (low fat)	Negative	No
	Diced breast and thigh	Negative	No
Poultry-duck samples	Breasts	Negative	No
	Breasts	Negative	No
Poultry-poussin	Whole	Negative	No
Beef	Mince	Negative	No
	Sirloin Steak	Negative	No
	Stewing steak	Negative	No
Lamb	Diced breast	Negative	No
	Mince	Negative	No
	Chops	Negative	No
Pork	Mince (low fat)	Cc (35)	Yes
	Chump steak	Negative	No
	Sausages –1	Negative	No
	Sausages –2	Cc (35)	No
	Sausages (frozen)	Negative	No

Negative samples were shown by a C_T value of 40

4.4 Discussion

4.4.1 Assay Development and Optimisation

The Taqman PCR assay described in this chapter was able to rapidly detect and definitely identify the major human enteropathogens *C. jejuni* and *C. coli*. It is an improvement over the conventional methods for speciation and has allowed substantial savings in time, enabled faster throughput and reduced the consumable costs to £0.84 (per isolate per Taqman reaction) compared to £1.80 by previous methods (biotyping). This was achieved by the optimisation of primer and probe concentrations to a minimum optimal value, by reduction in reaction volume to save on mastermix costs and with the additional use of cell lysates as a DNA isolation method to reduce cost and set up time without compromising specificity. The gene targets used with the primers and probes have been shown to be specific, since no amplification occurred with any of the other bacterial DNA tested from other *Campylobacter* species, related organisms or other enteric pathogens.

The primers and probes designed using Primer Express were shown to successfully amplify each target sequence. Design of the *map*A target primers and probe was straightforward with the use of a region of the gene, which met the design criteria in terms of GC content and also was specific for *C. jejuni* by testing in BLASTN. For the *ceu*E target an alignment was constructed to determine areas of divergence between the two genes (X88849 and X82427) where primers and a probe could be placed. The area chosen was shown to be successful for discrimination of *C. coli* from *C. jejuni* and this was validated with confirmatory testing. It could have been possible to use other areas of the *C. coli* sequence, as additional divergent areas were present, however the primers and probes selected worked well, therefore no further

areas were investigated. The differences in the *ceu*E gene and the related homologue in *C. jejuni* could have been exploited for the assay as an alternative to using the *map*A gene for *C. jejuni,* as has been described in the conventional PCR by Gonzalez *et al* (Gonzalez *et al.,* 1997). However the use of the different gene targets for the differentiation of each species ensured greater validity of results by the use of a *C. jejuni* specific gene (*map*A).

Optimisation of the reaction by testing variations in concentrations of the primers and probes was important to establish the best conditions for the PCR reaction. The lowest concentration of primers and probes was established (by C_T and ΔR_n) which provided optimal experimental data and maintained low reagent costs. This was particularly important for the concentration of probe, which was the most expensive reagent. Determination of the optimal probe concentrations for each target was useful in combining the two assays to create a duplex reaction. Both reactions used equimolar concentrations of primers and probe, and produced similar positive C_T values, therefore when the two assays were combined in equal concentrations the results were consistent, and no further testing was required.

Additionally, the reaction volume was reduced by half to lower reagent costs, where minimal effect on the C_T value was seen for each target gene. This meant substantial savings in mastermix and probe costs, when large numbers of samples were being processed. It would have been possible to reduce the reaction volume further to 12.5µl however the addition of a small quantity of DNA (2.5µl) was considered difficult and prone to errors although could have been achievable with the use of robotic equipment.

Sensitivity of the assay was tested to determine the lowest quantity of *C. jejuni* or *C. coli* DNA that was detectable. Sensitivity of the assay was confirmed with detection limits of 12.5 and 174 genome equivalents for *C. jejuni* and *C. coli* respectively per PCR reaction. The higher detection limit for *C. coli* could be explained by the absence of an accurate measurement for the size of the *C. coli* genome, due to the fact that the genome has not yet been sequenced. Only with the sequencing of the *C. coli* genome and a calculated genome size could an exact estimation of the sensitivity be made. Nevertheless, the assay has been shown to be specific and would be useful for detecting *C. jejuni* or *C. coli* DNA from food or environmental specimens where the quantity of DNA would be expected to be low. These results are comparable with another described Taqman assay for *C. jejuni* where a detection limit of 12 genome equivalents was calculated with a Taqman assay by Sails *et al* (Sails *et al.,* 2003a). Additionally other Taqman assays have been described for *C. jejuni* with detection limits of 1 CFU per PCR reaction (Nogva *et al.,* 2000a; Yang *et al.,* 2003). However the detection limits have been reported in CFU which are not comparable and also not as accurate as determining the sensitivity in genome equivalents. There are currently no described Taqman assays for *C. coli* against which to compare the results.

Optimal specificity of the assay was vital for its use as a valid detection method. This was tested with NCTC strains of all other species of campylobacter, related organisms such as *Helicobacter* and *Arcobacter* as well as other enteric organisms including *E.coli.* Results confirmed that the assay was specific for *C. jejuni* and *C. coli* and additionally the genetically distinct, but closely related subspecies of *C.*

jejuni, subspecies *doylei.* A problem reported with some of the conventional PCRs described previously (On & Jordan, 2003) was that the sequence variation in *C. jejuni* subsp *doylei*, had not been taken into account, therefore some of the PCRs were not detecting *doylei* strains. However this is not apparent with the assay described herein, where the NCTC *C. jejuni* subspecies *doylei* strain was identified as *C. jejuni.* The ability to detect this species provides an additional benefit, as subspecies *doylei* is rarely seen in human disease (On & Jordan, 2003), but could be present in environmental samples.

4.4.2 Comparison of DNA Extraction Methods

Three methods for DNA extraction, which were readily available, were evaluated with this assay. The MagNApure with two different kits, (i) the bacterial DNA Isolation kit and (ii) the Total Nucleic Acid isolation kit, which were compared against the crude cell lysate preparations. In this case, the limits of detection were determined from the input of CFU, where batches of the same concentrations of cells were used to test the different methods. The bacterial DNA isolation kit on the MagNApure resulted in superior levels of DNA recovery when compared to the other methods tested. Despite the superior performance of this kit on the MagNApure, for day-to-day applications sufficient DNA yield was achieved with the crude DNA preparations (cell lysates). The use of cell lysates allowed the assay set up to be a quick process only requiring a ten minute boiling step and without the additional time required for DNA extraction. This is only achievable due to the weak and fragile membrane of campylobacters, which is easily ruptured by heat treatment. A further benefit of using cell lysates is the reduced cost, of £0.05, a substantial saving when compared to the cost of utilising DNA extraction kits.

There are a few negative aspects to using the cell lysate method for DNA preparation, for example run to run variability may be an issue, and contamination may be cause for concern. Additionally, no assessment of the quality of the DNA isolated by the method was made. It would be expected that both methods on the MagNApure would result in a DNA yield of good quality and also be free of any interfering components. This type of extraction would be important if extensive DNA analysis was required, for example if DNA sequencing or fragment analyses were to be carried out, or if the DNA was intended for long term archiving. The crude cell lysate preparations were only suitable for short term testing, and did not store well for long periods due to the DNases released following heat treatment which degraded the DNA resulting in loss of template (Nogva *et al.*, 2000a). These components would usually be removed by a standard DNA extraction method. However the cell lysates were not required further following testing with the Taqman, and were therefore discarded.

An additional problem in the use of cell lysates may be the possibility of the DNA not being released. Mohran *et al* (1998) reported that up to 20% of campylobacter strains were resistant to lysis by boiling water and therefore further enzymatic stages or other DNA isolation methods may be required. The existence of two phenotypically distinct subgroups of campylobacter was speculated, based upon the ability to release, or not release PCR detectable DNA. This was described as being due to the blockage of high molecular weight molecules with extended conformations such as chromosomal DNA, by the smaller molecules (RNA and proteins), which can take on compact arrangements and pass out more easily

(Mohran *et al.,* 1998). This was not seen in the results described within this section, however based upon this study it would probably be advantageous to retest any suspect *C. jejuni/coli* isolates, which failed on the Taqman, but utilising MagNApure DNA isolation.

The MagNApure is a versatile labour saving platform capable of rapid DNA extraction from 32 samples and has been shown to be a flexible and reliable platform suitable for a variety of applications (Germer *et al.,* 2003; Knepp *et al.,* 2003). It provides a reliable and robust method for producing DNA of standardised quality and quantity if the same input of starting material is used. These findings are useful for the application to other PCR based assays. The superior performance of the MagNApure Bacterial DNA Isolation kit would lead this kit to be the method of choice to use in future investigations, especially where detection of small quantities of DNA would be important. However for extractions direct from culture then the use of the Total NA isolation kit could preferentially be used.

4.4.3 Large Scale use of the Assay

Large-scale use of the assay revealed the benefits for the speciation of samples in a timely manner. The number of isolates speciated within this six-month time frame could not have been achieved with conventional PCR or phenotypic speciation methods without significantly increased outlay of equipment and person hours. It has facilitated a new approach for the rapid speciation of campylobacters, applicable for work in high throughput laboratories.

The six months results have shown that 97.7% of the samples received in the laboratory were correctly identified as either *C. jejuni* or *C. coli*. Eleven strains (0.2%) were identified as non-C. *jejuni/coli* but as other *Campylobacter* species, (two *C. upsaliensis*, one *C. fetus* and eight *C. lari*), using 16S rRNA RFLP (Cardarelli-Leite *et al.*, 1996). However the isolates (2.1%), which were unidentified as other species, and by the Taqman is of concern.

The lack of specificity in the Taqman assay was addressed, in order to determine whether it was failure of the primers or probes, or the absence or variation in the genes, which was responsible for this small proportion of negative results. Initially it was considered to be due to a lack of PCR detectable DNA by boiling as reported by Mohran *et al* (1998). However for the sample set used in this chapter, the isolates which failed on the Taqman were subsequently identified by the use of the Vandamme primers in a conventional PCR using the same cell lysate, indicating that the lysate did contain released DNA. At the time of writing (27/9/03) very few studies had been described where one method for speciation had been applied to as many isolates (n=6015) as is reported here. It is plausible to speculate that the application of any PCR technique to this number of strains would have resulted in a similar number of strains failing. It is the nature of a variable organism such as campylobacter that no typing technique is going to be 100% successful. It is always the case that single nucleotide polymorphisms will occur in DNA as a result of insertions and deletions due to environmental stress and adaptation. It is noteworthy that many of the isolates, which failed were from an environmental source. Since the *ceuE* gene is postulated as being a potential virulence gene due to its involvement in iron sequestration, it is possible that divergence has occurred within the gene, away from the published sequence as a result of environmental influences.

4.4.4 Isolates that amplified both genes

Interestingly 15 strains (0.2%) amplified both genes. Possible explanations are that either the Taqman assay was non-specific; cells of both species were present i.e. the patient had a mixed infection or that these single strains possessed both genes. It was considered unlikely to be a result of the assay being non-specific, due to the small proportion of strains (0.2%) involved and the evidence from evaluation studies that the primers and probes were specific. When multiple single colony picks were taken and tested with the Taqman assay, ten out of fifteen were seen to be true mixed infections with both *C. jejuni* and *C. coli* strains. Single colony Taqman results were concordant with the sero/phage type species designation. It has been demonstrated that coinfection can occur with *Campylobacter* species (Richardson *et al.,* 2002) There is a growing amount of evidence that mixed infections of either species or serotype are significant features of Campylobacter infection and it would be interesting to use a modification of this technique to explore this phenomenon further.

However, in none of the remaining five strains where both genes were detected was it possible to separate the two species following multiple single colony picks. For these strains, results from sequencing the genes and other speciation tests showed no conclusive specific designation and interestingly all five were sero-untypeable. These data suggest there are a very small proportion of strains, which are "hybrids" with respect to the two genes used. As described earlier, it is well established that genetic heterogeneity in *C. jejuni* species can occur as a result of changes in environmental

conditions; therefore some of the mixed reactions may be a reflection of the natural genetic instability (Wassenaar *et al.*, 1998) of campylobacter.

It has been suggested that there are undescribed members of the genus campylobacter (Atabay *et al.*, 1997). This is exemplified by the identification of novel catalase negative, urease positive campylobacters which have been isolated from cattle faeces (Atabay *et al.*, 1997). The phylogenetic group, rRNA superfamily VI (Vandamme *et al.*, 1991) to which *C. jejuni* and *C. coli* are assigned exhibits a high degree of heterogeneity demonstrated by the diverse characteristics of its members. Since this classification, amendments have been made to take account of new species, or reclassification of existing ones. Therefore, in time it is highly likely that further *Campylobacter* species may be identified which are currently unseen due to current isolation practices. Hybrid strains such as these *mapA/ceuE* isolates may be representative of a further subset of species and further studies of the distribution of these strains would be valuable.

4.4.5 Advantages of speciation using the Taqman Assay

This assay enabled accurate speciation of 97.6% *C. jejuni* and *C. coli* strains tested. "Mixed strains" would previously have been unrecognised by phenotypic testing, as a positive hippurate assay would only have identified the *C. jejuni.* A proportion of strains failed in this assay, but these were identifiable as *C. jejuni* or *C. coli* with other methods and a few strains were identified as other *Campylobacter* species. Despite this, the throughput and high proportion speciated gave time and cost gains compared to phenotypic or conventional PCR methods.

Conventional PCR methods involve amplification of target genes followed by separation of the product on agarose gels. Sequence specific validation of PCR products is performed by time-consuming hybridisation analysis using blots or gels. Real time PCR using Taqman technology offers a rapid one-step PCR amplification and sequence specific validation. A greater throughput has been achieved than would have been possible using conventional PCR, where in comparison, assay completion time would be in excess of 4 hours (for a 96 well plate thermal cycler) plus time for gel electrophoresis. The closed tube of the Taqman system helps to eliminate contamination problems, reduces hands on time, and facilitates large scale sample processing.

Taqman PCR assays use small amplicons, up to 100bp, compared to the larger 200bp amplicon required for a standard PCR and separation on agarose gels. Furthermore, all Taqman assays are designed to amplify under the same cycling and reaction conditions making possible the running of more than one assay on one reaction plate. Further primer and probe combinations could be included in the same reaction plate allowing for identification of other *Campylobacter* species during the same run.

Separate positive and negative controls were included in every plate to verify that the amplification and reagents were working correctly. The inclusion of an internal positive control into the mastermix would be of even more benefit. For a single target assay a region of a distinct plasmid or sequence is used, which is amplified independently, leaving the main PCR cycling unaffected (Drosten *et al.*, 2001). This is labelled with a different fluorescent probe, which can be detected after the reaction to verify successful amplification. Even though more than two types of Taqman

probe are available, the limitations of the Taqman software are such that, only two fluorescent dyes can be detected per reaction, therefore the inclusion of a positive control was impossible in this assay. With improvements to the Taqman software the incorporation of a control may eventually be possible, which would provide further verification of assay results.

Reproducibility can be a problem issue with conventional PCRs. Despite the use of published primers, PCR reaction conditions and components, slight variations in reagent manufacturer, type of PCR machine and variations in running conditions can cause false negative results to occur (On & Jordan, 2003). Taqman based assays have the potential benefit that all assays developed are designed to run under standardised reaction conditions and with the same mastermix which makes assays comparable between different laboratories. With the large numbers of isolates tested here, reproducibility was demonstrated with similar C_T ranges obtained for positive samples and results consistent between runs.

4.4.6 Adaptation to the Roche Lightcycler

The Lightcycler results with both formats were consistent with the results from the Taqman, proving the assay was applicable for use on a different platform. This would be especially useful, for example, in the case of an urgent specimen whereby a single result for a single strain would be required quickly. In this case the Lightcycler would be more appropriate due to its fast cycling time of one hour. The Taqman is designed for the throughput of multiple samples and is more expensive if fewer than ten samples are tested in one run, due to the cost of the sample plates (£5 per plate). The lack of a probe in the Lightcycler SYBR Green assay makes this less sensitive

than the Taqman assay. The inclusion of sequence dependent probes in real time PCR take the assay up to another level of sensitivity compared to assays based on primers alone. This is due to the binding of three sequences of DNA in a Taqman reaction (or four as in the case of the Lightcycler using hybridisation probes). Nevertheless, the Lightcycler assay utilising the Taqman *mapA* and *ceuE* primers, is more specific than a conventional PCR due to the incorporation of SYBR green DNA binding dye, the small highly efficient amplicons which are produced and the subsequent melt curve analysis for verification of product size. It would require a large-scale study utilising the Taqman primers on the Lightcycler to determine the usefulness of the assay or the additional design of hybridisation probes for both gene sequences for the added sensitivity.

Taqman probes on the Lightcycler were determined to be successful with consistent results compared to the other two formats. Due to the constraints of the Lightcycler's detection channels, which cannot be altered, the fluorescent emissions from Taqman probes cannot be 100% accurate. Channel 1 on the Lightcycler detects emissions of wavelengths of 520nm, the Taqman probes emit fluorescence at wavelengths of 510nm (FAM) and 560nm (VIC), and therefore only the shoulder of fluorescence emissions is likely to be detected by the 540nm channel on the Lightcycler. This assay was only tested on 30 samples, so the likelihood of false negative results may not have been thoroughly investigated. This would be especially important if the assay were to be used for samples with low quantities of DNA for example in testing of food samples. Using Taqman probes on the Lightcycler also means the ability for melt curve analysis is lost, which normally on the Lightcycler provides a useful check for product size verification. The redesign of the probes as hybridisation

probes but covering the same gene regions as described, would make the assay more reliable and robust for the Lightcycler system. The Lightcycler Model 2 has recently become available. This offers the capability of using Taqman probes as well as up to 4 sets of Lightcycler hybridisation probes. This new instrument would allow for the redesign of hybridisation probes for each of the *mapA* and *ceuE* gene regions with the addition of an internal positive control. The new instrument still only processes 32 samples in one run; therefore the Taqman still remains the best machine for high throughput studies. Alternatively the Taqman probes could be used on other real time platforms including the Smartcycler (Cepheid, USA) or the icycler (Biorad, UK) which have more detection channels, therefore the capability of using hydrolysis probes efficiently, however both these systems are limited in the maximum number of samples which can be amplified in one run.

4.4.7 Direct detection studies

4.4.7.1 Direct application to charcoal transport swabs

The direct isolation of campylobacter genomic DNA from charcoal transport swabs using the MagNApure followed by successful speciation of samples with the Taqman assay was evaluated. Charcoal transport swabs are used routinely for the easy and safe transfer of cultures between laboratories. The direct extraction of DNA from the culture medium itself allows investigation of samples on receipt without waiting for culture, when time is limited. This strategy has reduced the time for direct speciation of *C. jejuni* and *C. coli* to between 4-5 hours depending on the number of samples, making speciation easily achievable within the working day. Moreover, as pure and high quality DNA is isolated, it would be suitable for use in further molecular

characterisation methods if required, for example sequence based typing such as MLST.

When using a direct extraction from the charcoal transport swab assumptions would be made that the organisms were still viable, and the swab would be cultured as normal to provide confirmatory testing. It would be unlikely for a transport swab to contain non-viable cells, as the swabs are designed to maintain the organisms within optimal conditions and these are usually cultured and incubated on receipt. Additionally a large amount of culture is usually collected onto the swab to allow for some cell loss during transport. If culture of a swab produced no detectable growth then this method had the advantage that it provided identification of the DNA within the cells. Failure to grow a campylobacter only occurs in approximately 1% of all samples received by the reference laboratory. The main benefit of the method is quick species testing, particularly in outbreak investigations, or when linking a suspect source to a patient sample.

4.4.7.2 Direct Application to artificially and naturally contaminated food samples

(i) Artificially Contaminated Food Samples

DNA extraction methods capable of detecting low levels of bacteria in complex matrices such as food samples are critical for successful PCR detection methodologies. Three different methods for DNA extraction from a range of different artificially contaminated food samples were tested and compared by the number of *C. jejuni* cells recovered. The MagNApure utilizing the Bacterial DNA extraction kit showed the highest DNA recovery rate in terms of the lowest number

of CFU still giving a positive C_T of 39 or lower. [The MagNApure in conjunction with this kit as already described is a method which is simple to set up rapid provides high recovery rates from non-cultured samples and additionally is a walk away system allowing time for other activities whilst DNA is being isolated]. Despite the success of the machine for DNA extraction it is an expensive piece of highly specialized laboratory equipment, which is. However it has additional capability of post elution set up for Taqman or Lightcycler capillaries, providing a potentially contamination free system for PCR set up.

The Bugs 'n Beads kit provides an easy to use method, which is based upon use of paramagnetic bacteria binding beads, followed by lysis steps, washing steps and elution of the bead-DNA complex. The only piece of ancillary equipment required is a magnetic separator, which is relatively inexpensive (£200). The generic method was straightforward and worked well for all the sample types investigated. Additionally, the procedure is amenable to automation, important for large-scale screening. Detection limits for milk samples of 1000 CFU/ml were obtained which were within ranges reported by others using this method, for example Nogva *et al* (2000b) reported a detection limit of 200 to 2000 CFU /ml of *Listeria monocytogenes* from skimmed and unpasteurised milk.

One bonus factor of the technique was that one of the first steps involved the complete removal of the food sample, leaving only the beads on the side of the tube. For this reason this method was expected to provide very good results, as any interfering factors within the food should have been removed in the first step, however the recovery was lower than that achieved with the MagNApure, but equal

to that of the Prepman system. It is possible that the addition of the bead-DNA suspension into the Taqman reaction adversely affected the Taqman chemistry, especially the fluorescence readings. The beads would settle to the bottom of the well and not remain dispersed throughout the solution, therefore suggesting that the readings taken were slightly inaccurate due to interfering components. Another factor was that the beads had to be resuspended well before the addition into the Taqman reaction, this was especially important when the dilution series was carried out. Sometimes the binding between the magnet and the beads was not optimal and some of the beads were lost as the liquid was removed from the tubes, suggesting that reproducibility was not the best possible.

This method has also been used in conjunction with a Taqman assay by Nogva *et al* (2000b) where they reported the Bugs 'n beads method to be superior to other magnetic bead based methods, however reproducibility problems were an issue. Since the beads are easy to manipulate in automated systems, have a microscopic size and the reduction of the Taqman reaction to nanolitre scale (Kalinina *et al.,* 1997), then paramagnetic beads and real time PCR strategies may represent tools for future detection systems.

The Prepman Ultra method was the easiest method to use, due to the one reagent formulation. This quick method was easy to apply and required no ancillary equipment apart from a standard laboratory centrifuge and heating block/water bath. This method worked well for all the samples tested and gave detection limits similar to those for the Bugs 'n Beads system. Its efficiency has been described for PCR and PCR-ELISA techniques from cultured samples of *E. coli* and *C. parvum* (Higgins *et*

al., 2001), from enrichment broths (Bolton *et al.*, 2002) and directly from ground beef where 10^5 cells were detectable per gram (Ge *et al.*, 2002). However the technique did not seem as robust as the Bugs 'n Beads system for some of the food samples tested, in particular the samples containing excess fat, including the milk and duck. In these samples, following the final centrifugation step a layer of fat remained on top of the supernatant, which led to difficulties in removing the clear (DNA containing) phase below. The pipette tip had to be placed through the fat layer into the clear layer below, a portion carefully removed and aliquoted into a sterile vessel. Failure of the sample would result if any of the fat layer was added into the Taqman reaction, for these reasons the Prepman system was not suitable for automation.

(ii) Naturally contaminated food samples

Limited methods exist for the direct detection and speciation of campylobacters form food matrices therefore a rapid, sensitive and specific method for the detection of *C. jejuni* and *C. coli* has been developed in union with a sensitive method for DNA extraction. This can be used on potential sources of infection and offers potential benefits for timely *C. jejuni* and *C. coli* detection.

There are problems when it comes to testing of *C. jejuni* in naturally contaminated food samples when using culture methods, as is the case for the detection of all food borne pathogens. This is due to the small numbers of cells present, the lack of methods for separation of pathogens from sample particulates, the presence of inhibitory compounds and additionally for *C. jejuni* the slow growth rate of the organism. All current methods employing cultural techniques are time consuming,

requiring prolonged incubation, and multiple subcultures, which are often not prolifiC. Additionally the cells may enter a non-viable state due to physical stress or starvation making them undetectable by culture. In this study a real time PCR assay has been applied to specifically detect *C. jejuni* and *C. coli* from potential naturally contaminated meat samples and the results compared with those from culture. The primers and probe set were shown to be specific for *C. jejuni* or *C. coli*, avoiding any artefacts in a potentially mixed population due to competition for the primers by amplification of spurious targets from other bacteria.

The results from the Taqman supported the fact that retail meat is commonly contaminated with campylobacter. All the samples, which were positive by culture, were also positive by Taqman, however 7 samples, which were negative by culture, were positive by Taqman. These positive results not confirmed by culture may be samples which contained dead or damaged cells which were at a sufficiently high concentration for them to be detectable by Taqman PCR however these could not be recovered by enrichment or culture. This may explain the failure of cultural techniques to detect organisms from potentially contaminated food samples and substantiates the claim that traditional culture methods result in significant undereporting of potentially infectious *C. jejuni* (Nogva *et al.*, 2000a) or *C. coli*. The concept of undetectable campylobacters by culture has been exemplified in transmission studies where no detectable growth by culture has been found (Atabay & Corry, 1997), but damaged or stressed cells or those in the reported VNC state have been described (Pearson *et al.*, 1993). Moreover it has been described that campylobacters in the described "VNC" state are more resistant to food processing

techniques than can be cultured, emphasizing that there is a requirement for valid methods of detecting campylobacter DNA from food samples (Pearson *et al.,* 1993).

Many different procedures have been described for the isolation of campylobacters from food samples, yet no unanimous method has yet been developed that is appropriate for all sample types and the choice usually depends upon the probable level of campylobacter within the sample material and any superfluous bacterial cells that may be present. One of the main drawbacks of direct extraction techniques either from commonly used matrices such as transport swabs or food samples is that there is no indication of the viability of the cells contained within the sample. This is one of the main challenges for DNA-based methods in diagnostics. It is generally assumed that the genomic DNA from bacteria persists after the organism dies, therefore the presence of DNA is not a good indicator of the viability of the organism (McKillip *et al.,* 1999). Alternatively RNA has been used as a marker for viable cells (Sails *et al.,* 1998; Sheridan *et al.,* 1998), however several assumptions have to be made based on the gene to be used. The gene has to be continually expressed, with an unstable transcript and a specific region has to be identified in the targeted gene, these difficulties make DNA a more preferable choice for use in viability studies. Acid-binding dyes which are reported to selectively enter dead bacteria with compromised membranes but do not enter viable cells due to the intact membrane systems and pumps (Rudi *et al.,* 2002), have been described as a potential strategy for distinguishing dead from viable cells. Additionally, certain nucleic acid binding dyes such as ethidium monoazide (EMA) can enter dead bacteria and be covalently linked to the DNA by photoactivation and therefore inhibit PCR. This was described by Nogva *et al* (2003) where a signal reduction by Taqman was seen for the treated

cells when compared to the untreated cells, representing a promising method for differentiation between viable/non-viable cells.

This assay demonstrated a good linear correlation between the amount of template DNA and amplified product (represented by C_T value) suggesting that the assay, although designed for end point detection, could be used for quantitative measurements. This would be useful for estimating the levels of campylobacters and therefore indicate potential risks.

4.4.8 Final conclusions

If we are to continue to understand the epidemiology of campylobacter infections and to be able to develop effective intervention strategies, the methods for detection and identification must be optimised for the purpose. As a molecular detection method for investigation of potential sources of campylobacter infection the assay developed in this study has provided a tool, which can be applied to the detection of a specific species within a suspect food sample. The assay is simple to perform, provides unambiguous results and can be completed within 3 hours (including assay set up time, running time and data analysis). However, no identification at the strain level other than the species is possible without substantial modification to the technique. The next sections of this thesis describe a method for further identification to strain level, whereby specific strains can be characterised to sub species level, allowing the detection of specific strains.

In conclusion, this assay performed on the Taqman has provided a rapid and reliable direct identification assay for *C. jejuni* and *C. coli*, which can be applied to crude cell

lysates of pure cultures and is now being used as a primary speciation tool in the CRU laboratory. This technique has also found wider application and is being used for speciation in the Manchester Health Protection Agency, Molecular Epidemiology Department, for isolates from human disease and the environment and has been used by Amar *et al* (unpublished data) within the Food Pathogen Reference Unit, Health Protection Agency London UK for identification of campylobacter DNA in stored faecal specimens. Additionally the adapted technique on the Lightcycler platform has been used by K Stapleton at the Veterinary Laboratories Agency, Weybridge, UK (Stapleton *et al.*, 2003).

Chapter 5

An assessment of the predictive alleles within MLST clonal complexes, for the use in strain detection and characterisation.

Chapter 5

An assessment of the predictive alleles within MLST clonal complexes, for the use in strain detection and characterisation of *C. jejuni.*

Multi Locus Sequence Typing (MLST) as described by Dingle (Dingle *et al.,* 2001) is rapidly becoming accepted as a widely applicable technique for accurate strain identification for *C. jejuni.* This has the advantages of a methodology that provides a discriminatory molecular profile, is reproducible with simple interpretation (Taylor & Fisher, 2003) and provides data which are directly comparable between laboratories via the website (http://pubmlst.org/campylobacter). Major genetic lineages or clonal complexes have been identified in *C. jejuni* populations from human infections, animal and environmental sources. One drawback of the system is that it can take up to three days to complete the PCR, sequencing and data analysis. The rapid identification of *C. jejuni* isolates to strain level, or initially clonal complex level would both inform and enhance the epidemiological investigation of *C. jejuni.* This would also provide a strategy whereby specific strains could be identified within environmental or food samples and subsequently be monitored throughout the food chain.

The next four chapters describe a novel strategy for the development of rapid real time allelic discrimination assays to identify strain associated single nucleotide polymorphisms (SNPs) based upon MLST locus alleles, which potentially identify clonal complexes. Chapter 5 describes the assessment of predictive locus alleles for use in identifying the six major clonal complexes ST-21, ST-45, ST-48, ST-61 ST-206 and ST-257. Chapter 6 investigates the application of detecting the predictive

locus alleles (identified in chapter 5) on the basis of SNPs. Subsequently; chapters 7 and 8 describe the design and development of allelic discrimination assays based on two different platforms for the rapid detection of SNPs. Additional investigation determined the applicability of using the assays directly to food samples, to provide rapid real time strain characterisation.

5.1 Introduction

At the time of writing (30/09/03) there was a substantial amount of sequence type information contained within the web accessible *C. jejuni* PubMLST database (http://pubmlst.org/campylobacter). This comprised the data from 2162 submitted isolates with 818 unique Sequence Types identified. Studies so far had shown that *C. jejuni* had a relatively small pool of unique alleles with few polymorphic nucleotides, although it was well established that these were stable polymorphisms, which identified the allelic profile of the various epidemiological lineages (clonal complexes) (Dingle *et al.*, 2001). The recognition of these major genetic lineages or clonal complexes in *C. jejuni* populations from human infections, animal and environmental sources (Dingle *et al.*, 2002) has provided major conceptual advances in our understanding of the population biology of *C. jejuni.*

The MLST clonal complex has been established as an important epidemiological group at the strain level providing accurate and phylogenetically valid strain identification, and enabling more amenable epidemiological analysis (Urwin & Maiden, 2003). Currently, twenty-four clonal complexes have been described (http://pubmlst.org/campylobacter), with ST-21 clonal complex being the largest, containing 26% of submitted isolates, and 60% of all isolates assigned to one of six major complexes these being, ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257. The data from preliminary MLST studies of *C. jejuni* isolates from animals and human cases of infection has established the concept of host associations or niche adaptation between clonal complexes of *C. jejuni* such as poultry, and cattle, which also cause human infections. For example, clonal complexes ST-45 and ST-257 are reported to contain isolates predominantly of poultry and human origin and ST-61 has been

associated with bovine and ovine isolates (Colles *et al.*, 2003; Dingle *et al.*, 2002; Manning *et al.*, 2003).

Based upon the intelligence that stable single nucleotide polymorphisms (SNPs) were present within the MLST alleles, and that the clonal complexes had been recognised as important strain groups, a strategy was proposed enabling the identification of specific clonal complexes by detection of SNPs. This chapter describes the first stage of a strategy for the identification of specific MLST clonal complexes by a SNP approach using allelic discrimination assays. The first step of which was the identification of the predictive alleles within six major clonal complexes (Table 5.1). The term "predictive allele" was used to define alleles, which were identified within a clonal complex, which were highly abundant within all the isolates currently assigned into the complex. e.g. the allele occurred within the majority of all the isolates within the clonal complex. Consequently, by identification of this predictive allele the clonal complex could be assumed. To determine the alleles, extensive investigation of the relationships between different allele types and clonal complexes was carried out by use of available MLST data currently in the PubMLST database.

The six target clonal complexes were chosen primarily due to being the major ones described (Dingle *et al.*, 2001; Dingle *et al.*, 2002) and were also accountable for 60% of human disease isolates in the largest MLST study published to date (Dingle *et al.*, 2001). Secondly, they were chosen due to the interesting host specific associations, which had previously been identified (Colles *et al.*, 2003; Dingle *et al.*, 2002; Manning *et al.*, 2003) (Table 5.1). Moreover, with the ultimate aim of

designing rapid allelic discrimination assays to identify the clonal complexes. These six clonal complexes were considered potentially valuable for epidemiological studies and for the possible identification of 60% of specific *C. jejuni* types (http://pubmlst.org/campylobacter).

Table 5.1 The six clonal complexes used, contribution to isolates on the database and host associations identified (Colles *et al.*, 2003; Dingle *et al.*, 2002; Manning *et al.*, 2003).

Clonal complex	Isolates within the database (%)	Host associations
ST-21	26.6	Diverse, especially human
ST-45	3.8	Poultry, cattle
ST-48	6.2	Widely distributed
ST-61	7.5	Cattle, sheep
ST-206	8.4	Widely distributed
ST-257	4.4	Poultry, cattle

5.2 Materials and Methods

Identification of the most common alleles contained within each of the six major clonal complexes (ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257).

Extensive analyses using the search functions of the *C. jejuni* MLST database were made, to determine the most common alleles within each clonal complex (http://pubmlst.net/campylobacter). Every clonal complex was investigated individually to determine the most abundant alleles at each locus. Using the "search database –advanced queries" function with the "search by clonal complex" box checked, allele distributions for each clonal complex were determined. Searches within each clonal complex, displayed the distribution of alleles across all the loci and alleles were selected, which were most specific for the six target clonal complexes ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257. All predictive alleles identified for each of the above clonal complexes were verified for exclusivity by using the "search database-advanced queries" function with the id field set as the allele and allele number entered.

Strategy for identification of predictive alleles:

- **Identification of the most abundant locus alleles within each clonal complex**

The statistics in the database were used to calculate the prevalence of each locus allele within every clonal complex. The most highly represented allele was used at each locus, and a positive predictive value was calculated. Positive predictive values indicated the degree of specificity, (the more specific, the closer to 1).

- **Comparisons across all clonal complexes**

This data was complied into a table where the highest scoring allele at every loci could be compared for every clonal complex. This enabled the identity of abundant alleles, which only occurred within the six target clonal complexes.

- **Identification of predictive alleles for the six target clonal complexes**

Predictive alleles were chosen based upon two criteria;

(i) High specificity of the allele for the target clonal complex

(ii) The exclusivity of that abundant allele for one particular target clonal complex.

- **Determination of specificity for the complexes ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257.**

The specificities of the identified alleles for the required clonal complex were tested from the whole database and predictive alleles were chosen for eventual assay design.

5.3 Results and Discussion

5.3.1 Abundant alleles within each clonal complex

For every clonal complex on the PubMLST database (last updated 30/01/04) data relating to the distribution of most abundant locus alleles was compiled by clonal complex and indicated by positive predicted values (Table 5.2). This displayed the distribution of the highest scoring alleles across each clonal complex allowing the most specific locus alleles within a clonal complex to be determined. In all six target clonal complexes high scoring alleles were identified with predicted values of 0.75 or higher, which could potentially be used for assay design. These were investigated further, for example within clonal complex ST-21 the allele *asp*A2 showed good specificity of 0.9, in contrast *glt*A1 showed specificity of 0.5 therefore was discounted from any further analysis.

5.3.2 Distribution of abundant alleles across clonal complexes

Figures 5.1 a-g were constructed to show the abundant alleles at each locus (calculated above) against the number of clonal complexes in which they occurred. This indicated the complexes in which these abundant alleles were present and determined those unique to the six target clonal complexes. Results (Table 5.2) showed that there were highly specific abundant alleles within each of the six target clonal complexes, which did not occur as abundant alleles throughout the remainder. However also within some loci there were many commonly occurring high scoring alleles, for example the allele *asp*A1 was an abundant allele within clonal complexes, ST-61, ST-179, ST-22, ST-42, ST-658 ST-508 and ST-362, therefore not a suitable predictive allele for a specific complex. This was also the case for alleles *gln*A2, *glt*A2, *gly*A2 *gly*A4, *tkt*_3 and *unc*A6, which were all abundant alleles occurring in

six or more clonal complexes. Additionally *asp*A2 was another frequent allele occurring as the most abundant allele within four clonal complexes. This allele has been described to be obtained from diverse isolates and diverse clonal complexes, (from humans, sand, and poultry) (Dingle *et al.*, 2001) therefore despite the high specificity for ST-21 (0.9) it would clearly not make a good predictive allele. Interestingly, all the alleles in Table 5.2 occurring as abundant alleles more than once were low number alleles, the higher number alleles were of far lesser occurrence. This was due to some alleles predominating within the MLST data set, with the remainder present within fewer clonal complexes. As the allele numbers were assigned and named in order of recognition (Dingle *et al.*, 2001), then it is not surprising that the higher number alleles of which a few occur within the table e.g. *pgm*_89 (ST-460) are within the less common clonal complexes. For these, there are only a small number of entries on the database, due to the more recent description of the clonal complex.

Table 5.2 Most abundant alleles at each locus within all clonal complexes. (Shaded areas indicating predictive alleles for the six target clonal complexes. (as of 30/01/04)

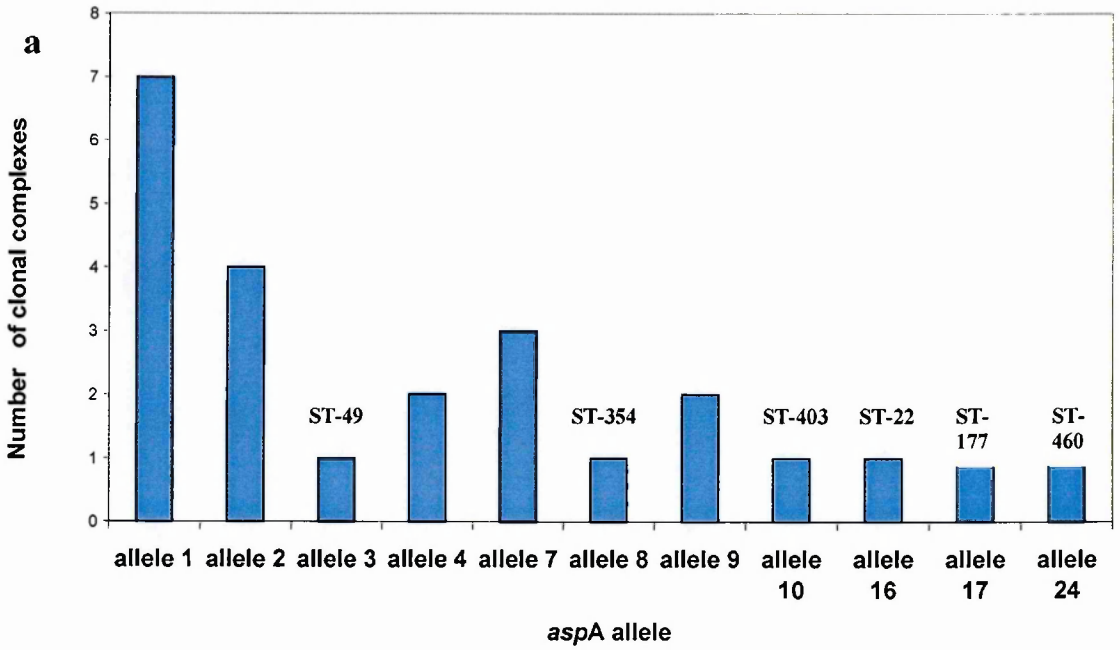
ST-Complex	21	257	61	48	45	206	403	177	179	22	49	52
(No. of isolates)	577	83	135	164	184	96	87	21	9	61	29	39
aspA	2 (0.9)	9 (0.76)	1 (0.93)	2 (0.78)	4 (0.84)	2 (0.82)	10 (1)	17 (0.95)	1 (1)	1 (0.91)	3 (0.93)	9 (0.7)
glnA	1 (0.93)	2 (0.9)	4 (0.95)	4 (0.94)	7 (0.96)	21 (0.72)	27 (0.87)	2 (0.57)	6 (0.89)	3 (0.9)	1 (0.93)	25 (0.64)
gltA	1 (0.53)	4 (0.92)	2 (0.96)	1 (0.68)	10 (0.88)	5 (0.96)	16 (0.75)	8 (0.8)	7 (0.55)	6 (0.91)	5 (1)	2 (0.94)
glyA	3 (0.90)	62 (0.85)	2 (0.94)	2 (0.81)	4 (0.9)	37 (0.56)	19 (0.94)	5 (0.95)	2 (0.89)	4 (1)	17 (0.86)	10 (0.94)
*pgm_*	2 (0.87)	4 (0.96)	6 (0.95)	7 (0.48)	1 (0.74)	2 (0.83)	10 (0.79)	8 (0.61)	40 (0.78)	3 (0.86)	11 (0.93)	22 (0.58)
tkt	1(0.92)	5 (0.96)	3 (0.97)	1 (0.76)	7 (0.87)	1 (0.96)	5 (0.78)	2 (0.9)	32 (0.78)	3 (0.95)	11 (0.96)	3 (1)
uncA	5 (0.82)	6 (0.93)	17 (0.76)	5 (0.98)	1 (0.95)	5 (0.97)	7 (1)	4 (0.6)	3 (0.89)	3 (0.93)	6 (1)	6 (0.87)

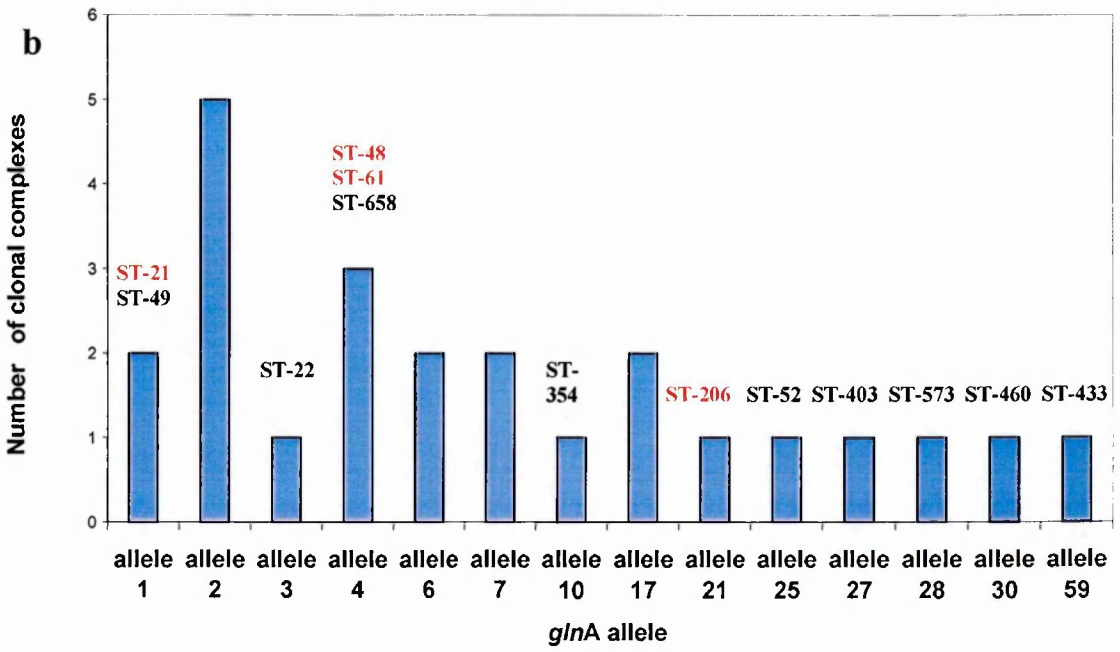
(Continued on next page)

Table 5.2cont

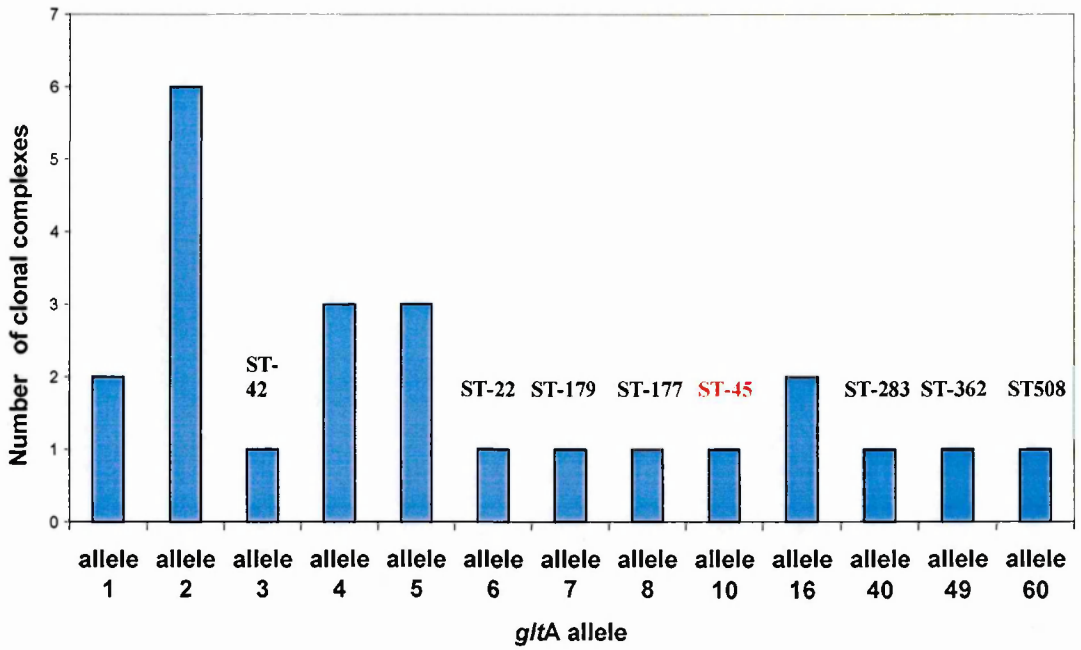
ST-Complex	353	354	42	283	658	460	573	508	443	41	362	433
(No. of isolates)	53	31	86	28	25	8	9	50	26	19	17	13
aspA	7 (0.75)	8 (0.67)	1 (0.96)	4 (0.6)	1 (0.24)	24 (0.62)	7 (0.7)	1 (1)	7 (0.5)	16 (1)	1 (1)	2 (0.76)
glnA	17 (0.69)	10 (0.64)	2 (0.96)	7 (0.96)	4 (0.8)	30 (0.62)	28 (0.88)	6 (1)	17 (0.84)	2 (1)	2 (1)	59 (0.69)
gltA	5 (0.84)	2 (1)	3 (0.89)	40 (0.7)	2 (0.92)	2 (1)	4 (0.88)	60 (0.92)	2 (0.92)	16 (1)	49 (0.82)	4 (1)
glyA	2 (0.92)	2 (0.83)	4 (0.95)	4 (0.75)	4 (0.8)	2 (1)	28 (1)	24 (1)	15 (0.88)	62 (1)	4 (0.89)	38 (0.84)
pgm	10 (0.56)	11 (0.83)	5 (0.90)	42 (1)	19 (0.56)	89 (0.87)	17 (0.6)	12 (1)	23 (0.88)	3 (1)	11 (0.94)	17 (0.92)
tkt	3 (0.9)	12 (0.9)	9 (0.91)	51 (0.78)	3 (0.96)	59 (0.75)	34 (0.88)	28 (0.98)	3 (0.80)	9 (1)	66 (0.82)	12 (0.92)
uncA	6 (0.9)	6 (0.96)	3 (0.82)	1 (0.96)	6 (0.96)	6 (1)	12 (0.77)	1 (0.96)	12 (0.88)	8 (1)	8 (1)	35 (0.84)

Figures 5.1a-g Representation of the most abundant alleles (calculated from Table 5.2) at each locus and the number of clonal complexes in which they occur, unique abundant alleles and ones used in combination are indicated. (red-alleles used for assay design)

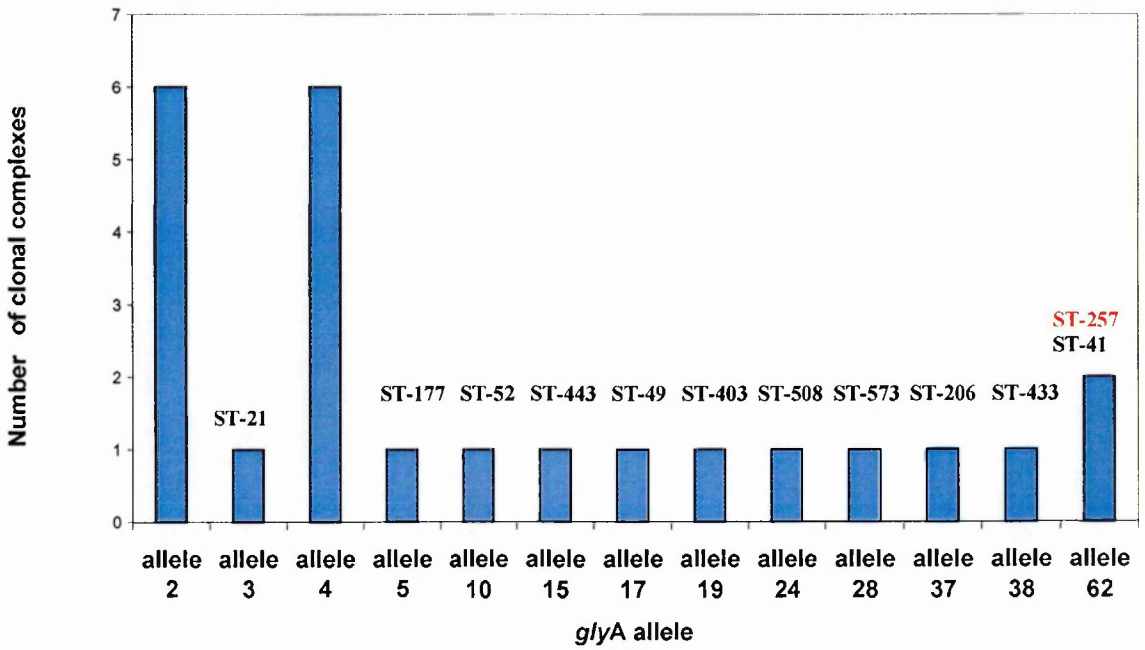




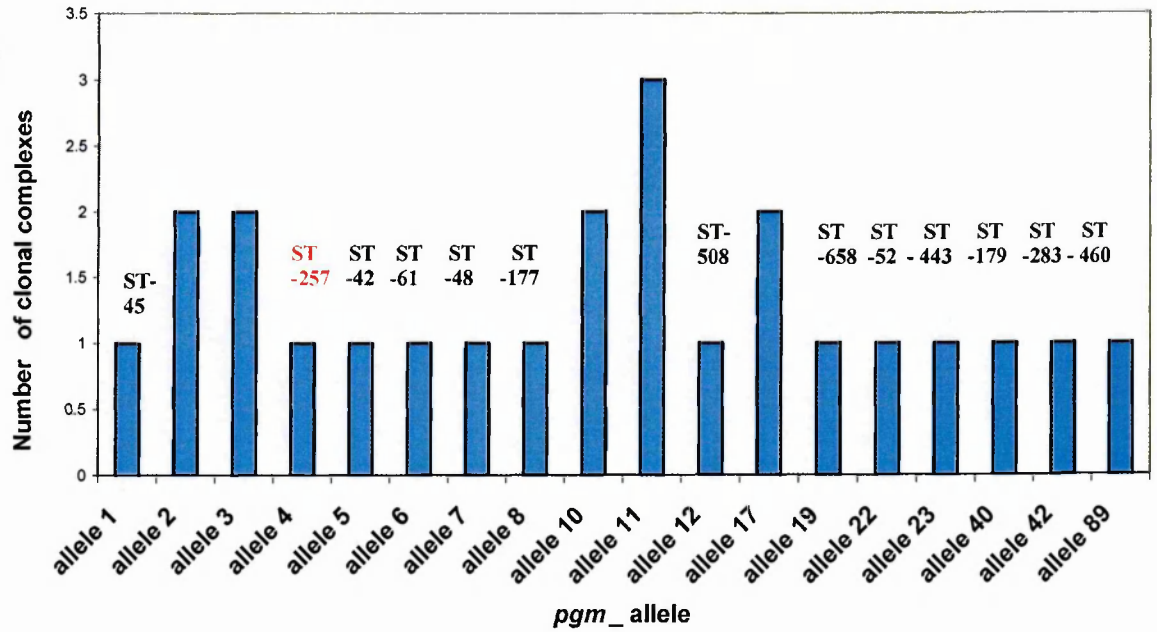
c


Number of clonal complexes
0
1
2
3
4
5
6
7
ST-42
ST-22
ST-179
ST-177
ST-45
ST-283
ST-362
ST508
allele 1
allele 2
allele 3
allele 4
allele 5
allele 6
allele 7
allele 8
allele 10
allele 16
allele 40
allele 49
allele 60
gltA allele

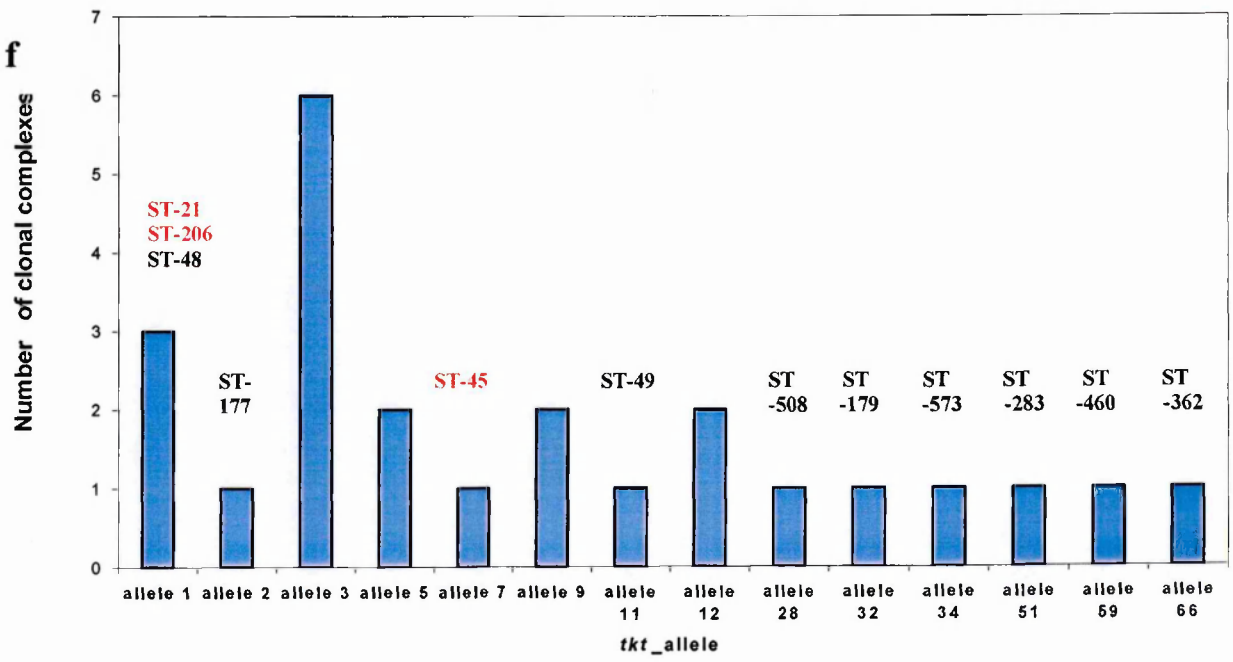
d


Number of clonal complexes
0
1
2
3
4
5
6
7
ST-21
ST-177
ST-52
ST-443
ST-49
ST-403
ST-508
ST-573
ST-206
ST-433
ST-257
ST-41
allele 2
allele 3
allele 4
allele 5
allele 10
allele 15
allele 17
allele 19
allele 24
allele 28
allele 37
allele 38
allele 62
glyA allele

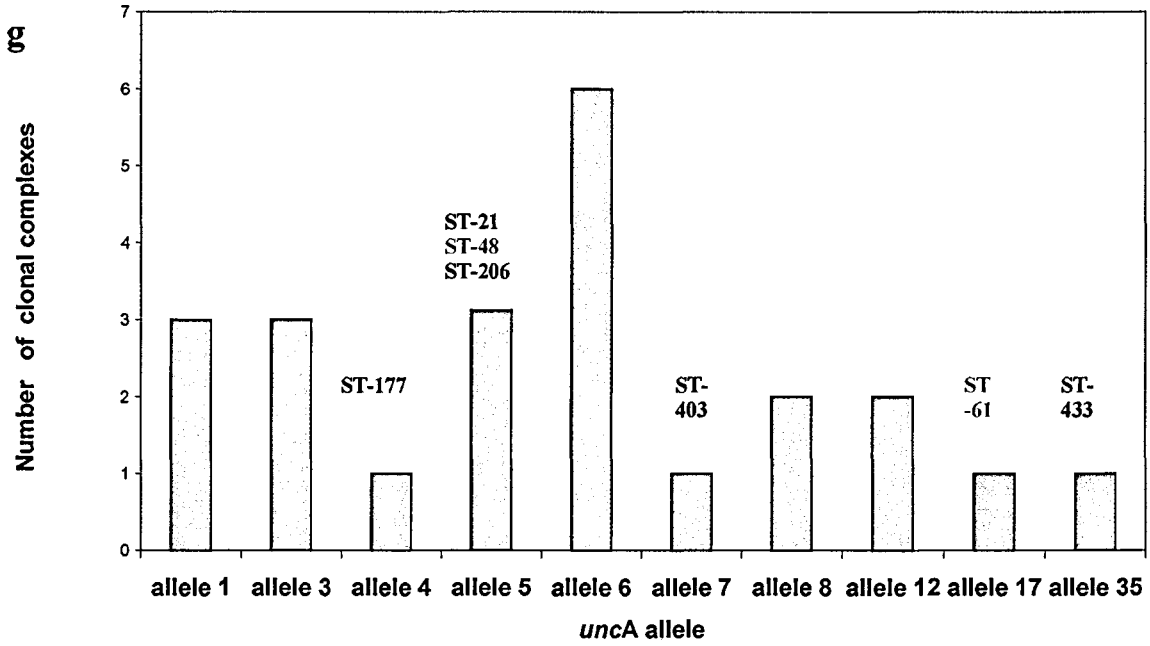
e


Number of clonal complexes
3.5
3
2.5
2
1.5
1
0.5
0
ST-45
ST-257
ST-42
ST-61
ST-48
ST-177
ST-508
ST-658
ST-52
ST-443
ST-179
ST-283
ST-460
allele 1
allele 2
allele 3
allele 4
allele 5
allele 6
allele 7
allele 8
allele 10
allele 11
allele 12
allele 17
allele 19
allele 22
allele 23
allele 40
allele 42
allele 89
pgm_ allele

f


Number of clonal complexes
7
6
5
4
3
2
1
0
ST-21
ST-206
ST-48
ST-177
ST-45
ST-49
ST-508
ST-179
ST-573
ST-283
ST-460
ST-362
allele 1
allele 2
allele 3
allele 5
allele 7
allele 9
allele 11
allele 12
allele 28
allele 32
allele 34
allele 51
allele 59
allele 66
tkt_allele

gg


Number of clonal complexes
7
6
5
4
3
2
1
0
ST-21
ST-48
ST-206
ST-177
ST-
403
ST
-61
ST-
433
allele 1
allele 3
allele 4
allele 5
allele 6
allele 7
allele 8
allele 12
allele 17
allele 35
uncA allele

5.3.3 Specific analysis of predictive alleles identified for clonal complex ST-21

Within the clonal complex ST-21 the alleles most abundant with high-calculated specificity (Table 5.2) were *glnA*1, *glyA*3 and *tkt*_1, indicating that these alleles could potentially be used in the assay design. Despite the other alleles *asp*A2 and *unc*A5 showing high specificities (0.9) they were also abundant in too many other clonal complexes and were discounted as potential predictive alleles. The remaining identified alleles *glt*A1 and *pgm*_10 were discounted from any further investigations due to their low specificities (0.5 and 0.56) (Table 5.3).

One of the most abundant alleles within this clonal complex was allele *glnA*1, which had a specificity of 0.93, however this allele also had high specificity for clonal complex ST-49 (0.9) (Table 5.3). From all the isolates on the database, the allele *gln*A1 showed high specificity of 0.9 for clonal complex ST-21 however it was also present in other clonal complexes ST-22, ST-48, ST-49, ST-52, ST-177, ST-206, ST-257 and ST-353, with lower specificities (0.002-0.04).

The allele *tkt_1* had a specificity of 0.94 within all ST-21 clonal complex assigned isolates; however this allele also had high specificity in two other clonal complexes ST-48 and ST-206 with specificities of 0.98 and 0.97 respectively. From the whole database the *tkt*_1 allele had a specificity of 0.67 for the clonal complex ST-21. This allele also occurred within other complexes in particular ST-48, ST-206 and ST-45, but with lower specificities (0.01-0.15). This allele was shown to not occur within clonal complex ST-49.

The allele *gly*A3 had a high specificity for ST-21 of 0.92 and additionally had no other high specificities within other clonal complexes. From the whole database this allele showed a specificity of 0.9 for ST-21, and low specificities (0.0018-0.01) for 6 other clonal complexes ST-206, ST-42, ST-45, ST-283, ST-48 and ST-61.

The alleles *gln*A1, *tkt*_1 and *gly*A3 were determined to be predictive alleles for ST-21 and could therefore be used either singly or in combination for the eventual assay design. It became clear throughout the investigation that especially for the clonal complex ST-21, more than one predictive allele would be required to achieve the required specificity. Therefore the two alleles *gln*A1 and *tkt*_1 were used together to identify the clonal complex ST-21, (also for reasons explained in more detail in the next chapter). The combined specificity for the two alleles together as predictive alleles for the clonal complex was calculated. It was determined that a combined specificity of 0.88 was obtained for the *gln*A1 and *tkt*_1 alleles for clonal complex ST-21 and a specificity of 0.98 was obtained for the identification of clonal complex ST-21 by use of these alleles from the whole database. These also had a small occurrence within the complexes ST-206 (0.004) and ST-48 (0.002); however this was overlooked for this assay design.

Low specificities of the alleles within other complexes were accommodated in this assay design by the use of two alleles for detection of the complex. Use of only the assay based on *glnA1* would have retrieved a significant proportion of ST-49 clonal complex assigned isolates and therefore would not be sufficiently discriminatory. Likewise, the specificity of the allele *tkt*_1 within the whole database was more frequent than *glnA1*, and it occurred in a number of other complexes.

The choice of alleles for the clonal complex ST-21 did not follow the original strategy, as both of the alleles used were not unique for ST-21 and therefore occurred within other clonal complexes. The process was more straightforward for the assay design for the other five target clonal complexes. This could be explained by the fact that ST-21 clonal complex is the largest, most diverse complex and also the most likely to be associated with human infection. The clonal complex ST-21 probably corresponds to a large stable cluster of isolates capable of colonising a wide range of hosts and may be well adapted for long term survival given the apparently ubiquitous distribution of isolates assigned into this complex (Colles *et al.*, 2003; Dingle *et al.*, 2002). For these reasons it was important to choose predictive alleles which allowed the accurate identification of the clonal complex with maximum specificity for ST-21. The alleles identified have been shown to be highly specific with a combined specificity of 0.98 for ST-21 clonal complex, confirming them as suitable for assay design.

Table 5.3 Associations of the predictive alleles used for identification of the ST-21 complex (shaded areas most applicable alleles and in combination)

Allele	Specificity of allele within ST-21 only	Other high specificity of allele (from table 5.2)	Specificity of allele from whole PubMLST database	
			Assigned to ST-21	Not assigned to ST-21
*asp*A2	0.9	ST-61 (0.78) ST-206 (0.82) ST-433 (0.76)	0.6	All complexes
*gln*A1	0.93	ST-49 (0.9)	0.9	ST-49 (0.04) ST-206,22,257,353,52 (0.003) ST-177,48 (0.002)
*glt*A1	0.5	ST-48 (0.68)	0.7	ST-48 (0.2) ST-42 443,61 (0.004) ST-206,353,52,573,658 (0.002)
*gly*A3	0.92	None	0.9	ST-206 (0.01) ST-42,45 (0.0036) ST-283, 48, 61 (0.0018)
*pgm*_10	0.56	ST-403 (0.78)	0.01	ST-403 (0.6) ST-353 (0.2) ST-22,354,433,52 (0.008)
*tkt*_1	0.94	ST-48 (0.75) ST-206 (0.96)	0.67	ST-48 (0.15) ST-206 (0.12) ST-45 (0.01) ST-257,61,42, 443,573,658 (0.003)
*unc*A5	0.9	ST-48 (0.98) ST-206 (0.97)	0.6	ST-48 (0.2) ST-206 (0.1) ST-61 (0.01) ST-45 (0.008) ST-443 (0.002) ST-177,22,353,42,433 (0.001)
Combined *gln*A1 *tkt*_1	0.88	None	0.98	ST-206 (0.004) ST-48 (0.002)

5.3.4 Specific analysis of predictive alleles identified for clonal complex ST-45

The alleles *gltA*10 and *tkt*_7 were identified as suitable predictive alleles for the clonal complex ST-45, which could possibly be used for assay design. Both these alleles had high-calculated specificities and were unique as abundant alleles for clonal complex ST-45. Other alleles *asp*A4, *gln*A7, *gly*A4, and *unc*A1 were identified with high specificities, however these showed high specificities also within other clonal complexes (Table 5.4), Additionally *pgm*_1 was shown to be unique for ST-45 but with lower specificity for ST-45 than *glt*A10 or *tkt*_7. All alleles except *glt*A10 and *tkt*_7 were excluded from any further investigations.

These alleles (*glt*A10 and *tkt*_7) were suitable for use as predictive alleles due to their high abundance within clonal complex ST-45 with specificities of 0.88 and 0.89 respectively and were both also unique as abundant alleles for this clonal complex (Table 5.4). Additionally both alleles *glt*A10 and *tkt*_7 had high specificity for ST-45 from the whole database (0.98 and 0.93 respectively). A small number of *gltA*10 alleles were assigned into other complexes (ST-21 and ST-283) but at low specificity (0.009). Also a small number of *tkt-7* alleles were assigned into other clonal complexes ST-283, ST-21, ST-353, ST-48, ST-508 but with low specificities (0.01-0.004).

Excellent specificity of 0.99 was seen when these alleles were used in combination to identify the clonal complex ST-45 from the whole database. This confirmed that the alleles used together in an assay would provide an accurate identification of any ST-45 assigned isolates. For this clonal complex the strategy for identifying predictive alleles was relatively straightforward due to the presence of alleles, which were both

unique to the complex as well as being very specific. Additionally, due to the high specificity of the allele *glt*A10 for ST-45 from the whole database it was considered that this allele could be used independently in an assay to define the complex. Then at a later date if greater specificity was required, then the allele *tkt*_7 could be incorporated.

Table 5.4 Specificities of alleles for characterisation of clonal complex ST-45 (shaded areas-most applicable alleles and in combination)

Allele	Specificity of allele within ST-45 complex only	Other high specificity of allele (from table 5.2)	Specificity of allele from whole PubMLST database	
			Assigned to ST-45	Not assigned to ST-45
*asp*A4	0.84	ST-283 (0.6)	0.8	ST-283 (0.06) ST-21 (0.03) ST-206,257 (0.01) ST-354,48 (0.007) ST-42,52 (0.0003)
*gln*A7	0.96	ST-353 (0.75) ST-283 (0.7)	0.87	ST-283 (0.1) ST-21,353,354 (0.004)
*glt*A10	0.88	None	0.98	ST-21 (0.009) ST-283 (0.009)
*gly*A4	0.8	ST-22 (1) ST-42 (0.95) ST-283 (0.75) ST-658 (0.8) ST-362 (0.89)	0.4	ST-42 (0.17) ST-22 (0.13) ST-48 (0.05) ST-283,658 (0.04) ST-362 (0.03) ST-403 (0.008) ST-21 (0.006) ST-61 (0.004)
*pgm*_1	0.76	None	0.9	ST-21 (0.02) ST-22,48,52 (0.005)
*tkt*_7	0.89	None	0.93	ST-283 (0.01) ST-21 (0.008) ST-353,48,508 (0.004)
*unc*A1	0.94	ST-283 (0.96) ST-508 (0.96)	0.6	ST-508 (0.12) ST-283 (0.07) ST-21 (0.04) ST-61,257 (0.01) ST-177,206,22,353,658 (0.003)
Combined *glt*A10 *tkt*_7	1	None	0.99	1 unassigned isolate

5.3.5 Specific analysis of predictive alleles identified for clonal complex ST-48

The two alleles identified as being potentially predictive for the clonal complex ST-48 were *gln*A4 and *unc*A5, which showed the greatest specificity within the ST-48 clonal complex. The remaining identified alleles *asp*A2, *glt*A1, *gly*A1, *pgm_7* and *tkt_1* showed lower specificities for this clonal complex and were therefore discounted from any further analysis (Table 5.5).

The *gln*A4 allele was suitable for use, due to having a specificity of 0.94 for ST-48, however this allele also showed high specificity within other clonal complexes ST-61 (0.95) and ST-658 (0.8). Due to the allele *gln*A4 occurring with high specificities within both clonal complexes ST-61 and ST-48, a combined approach for the identification of both target clonal complexes using the same initial predictive allele was considered. This meant that an assay to detect the presence of the allele *gln*A4 would precede specific separate assays for the predictive alleles for ST-48 and ST-61. A combined approach offered potential advantages in eventual assay set up time, reduced costs and made the assays more amenable to high throughput.

The allele *unc*A5 was determined as a suitable second allele for the subsequent identification of ST-48. This allele had high specificity of 0.98 for ST-48, however it also had high specificity for other clonal complexes ST-21 and ST-206. This allele showed a wide distribution and was present as an abundant allele within 3 other clonal complexes (Figure 5.1g), therefore it had a low specificity from the whole database of 0.2 and therefore could not be used as a predictive allele for ST-48 independently. A different allele could have been used, but other abundant alleles within ST-48 had lower specificity as well as high abundance within other

complexes. By using a combination of the alleles *gln*A4 and *unc*A5 for identification of ST-48 a combined specificity of 1 was determined within ST-48 and specificity of 0.8 from the whole database. There were only very low occurrences of these two alleles together in other clonal complexes ST-206 (0.1) ST-61 (0.04) and ST-21 (0.01). With assays being designed for these three clonal complexes then any false positive ST-48 would be detectable by additional assays.

Table 5.5 Specificities of alleles for characterisation of clonal complex ST-48 (shaded areas-most applicable alleles and in combination)

Allele	Specificity of allele within ST-48 complex assigned isolates	Other high specificity of allele (from table 5.2)	Specificity of allele from whole PubMLST database	
			Assigned to ST-48	Not assigned to ST-48
*asp*A2	0.78	ST-21 (0.9) ST-206 (0.82) ST-433 (0.76)	0.15	ST-21 (0.62) ST-206 (0.09) ST-45 (0.01) ST-433 (0.01) All other complexes (>0.009)
*gln*A4	0.94	ST-61 (0.95) ST-658 (0.8)	0.4	ST-61 (0.3) ST-206,658 (0.05) ST-21,354 (0.008 ST-22,257,353,45 (0.002)
*glt*A1	0.68	ST-21 0.53)	0.2	ST-21 (0.7) ST-42 443,61 (0.004) ST-206,353,52,573,658 (0.002)
*gly*A2	0.81	ST-61 (0.94) ST-179 (0.89) ST-353 (0.92) ST-354 (0.83) ST-460 (1)	0.4	ST-48 (0.4) ST-52 (0.1) ST-354 (0.09) ST-48,443,658 (0.07) ST-21 (0.04) ST-460 (0.02) ST-353,45,257 (0.006)
*pgm*_7	0.48	None	0.7	ST-21 (0.2) ST-179,206,45 (0.008)
*tkt*_1	0.76	ST-21 (0.92) ST-206 (0.96)	0.15	ST-21 (0.67) ST-206 (0.12) ST-45 (0.01) ST-257,61,42, 443,573,658 (0.003)
*unc*A5	0.98	ST-21 (0.82) ST-206 (0.97)	0.2	ST-21 (0.6) ST-206 (0.1) ST-61 (0.01) ST-45 (0.008) ST-443 (0.002) ST-177,22,353,42,433 (0.001)
Combined *gln*A4 *unc*A5	1	None	0.8	ST-206 (0.1) ST-61 (0.04) ST-21 (0.01)

5.3.6 Specific analysis of predictive alleles identified for clonal complex ST-61

For the clonal complex ST-61, all the identified alleles were unique in that they all showed high specificities (>0.90). However, many occurred as abundant alleles within other clonal complexes in particular *asp*A1, *glt*A2, *gly*A2, *tkt*_3 which were all abundant alleles within six or more clonal complexes. The only two identified alleles, which were not present within other clonal complexes, were *pgm*_6 and *unc*A17, which in addition to *gln*A4 were investigated further (Table 5.6)

The allele *unc*A17 was identified as a potential predictive allele for the clonal complex ST-61. Specificity was confirmed within the complex (0.93) and also from the entire database with a specificity of 0.99 for *unc*A17 occurring within ST-61. *UncA*17 was unique for the clonal complex ST-61. This allele had no other high occurrences calculated and also had little specificity within other complexes, only 0.009 for ST-21. This allele has been described as possibly originating from *C.coli* (Dingle *et al.*, 2002). The allele *pgm*_6 was also identified which had high specificity and was not abundant within other clonal complexes, however its specificity from the whole database for ST-61 was not as high as that of *unc*A17, therefore *unc*A17 was preferentially used. The specificity of *gln*A4 had been confirmed for ST-61 (0.95). This also occurred within other complexes, in particular ST-48 (0.94) as already described. However, by combining the two alleles *gln*A4 and *unc*A17 high specificity was achieved of 1 with this combination of alleles not occurring in other clonal complexes within the database.

Table 5.6 Specificities of alleles for characterisation of clonal complex ST-61 (shaded areas-most applicable alleles and in combination)

Allele	Specificity of allele within ST-61 only	Other high specificity of allele (from table 5.2)	Specificity of allele from whole PubMLST database	
			Assigned to ST-61	Not assigned to ST-61
***asp*A1**	0.93	ST-179 (1) ST-22 (0.91) ST-42 (0.96) ST-658 (0.24) ST-508 (1) ST-362 (1)	0.3	ST-42 (0.2) ST-22,508 (0.1) ST-362 (0.04) ST-179, 658 (0.02) ST-45 (0.01) ST-206,21,257 (0.005) ST-443,48,573 (0.002)
***gln*A4**	0.95	ST-48 (0.94) ST-658 (0.8)	0.3	ST-48 (0.4) ST-206,658 (0.05) ST-21,354 (0.008 ST-22,257,353,45 (0.002)
***glt*A2**	0.92	ST-52 (0.94) ST-354 (1) ST-658 (0.92) ST-460 (1) ST-443 (0.92)	0.4	ST-61 (0.4) ST-52 (0.1) ST-354 (0.09) ST-48,443,658 (0.07) ST-21 (0.04) ST-460 (0.02) ST-353,45,257 (0.006)
***gly*A2**	0.85	ST-48 (0.8) ST-179 (0.89) ST-353 (0.92) ST-354 (0.8) ST-460 (1)	0.3	ST-48,61 (0.3) ST-353 (0.1) ST-206,354 (0.06) ST-179,460,21 (0.01) ST-658,21 (0.007)
***pgm*_6**	0.96	none	0.7	ST-48 (0.1) ST-403 (0.08) ST-21 (0.9) ST-177,443 (0.006)
***tkt*_3**	0.96	ST-22 (0.95) ST-353 (0.9) ST-658 (0.96) ST-443 (0.8)	0.3	ST-22 (0.2) ST-353,52 (0.1) ST-658,443 (0.05) ST-21 (0.01) ST-283 (0.008) ST-354,403,42,460,48 (0.003)
***unc*A17**	0.93	none	0.99	ST-21 (0.009)
combined *gln*A4 *unc*A17	1	none	1	none

5.3.7 Specific analysis of predictive alleles identified for clonal complex ST-206

Within clonal complex ST-206 two potential predictive alleles were identified these being *gln*A21 and *tkt*_1. Other alleles with high specificities could have been used for example *asp*A2, *glt*A5, *pgm*_2 or *unc*A5 however these all showed high abundance within other clonal complexes and were therefore unsuitable for use. The other allele identified was *gly*A37, which was only abundant within ST-206, however it had a low specificity of 0.56 and was therefore not investigated further (Table 5.7).

*Gln*A21 was determined to have a specificity of 0.72 within the clonal complex ST-206, and it showed no other high specificity within other clonal complexes (Table 5.7). The specificity was also determined from the whole dataset where a specificity of 0.8 was determined for isolates with the *gln*A21 allele, which were assigned into clonal complex ST-206. Furthermore, this allele was potentially useful as it only had very low specificity for other clonal complexes.

The additional allele *tkt*_1 was incorporated to ultimately improve the specificity. This allele was the same as that used for identification of ST-21 clonal complex, however the specificity for clonal complex ST-206 of 0.96 was sufficient for it to be used as an additional predictive allele for ST-206. As in the case of the clonal complexes ST-61 and ST-48 a shared allele was used to ultimately aid the eventual assay design.

The data was used to calculate the combined specificity if the two alleles were used together to be predictive of the clonal complex. A specificity of 1 was obtained for the presence of both alleles within the clonal complex and these two alleles were not

shown to occur together in any other clonal complexes. The combined specificity of the two alleles together was 0.96 from the whole database; together these alleles only occurred within clonal complex ST-21 at very low specificity (0.04).

Table 5.7 Specificities of alleles for characterisation of clonal complex ST-206 (shaded areas- most applicable alleles and in combination)

Allele	Specificity of allele within ST-206 only	Other high specificity of allele (from table 5.2)	Specificity of allele from whole PubMLST database	
			Assigned to ST-206	Not assigned to ST-206
*asp*A2	0.82	ST-21 (0.9) ST-48 (0.78)	0.09	ST-21 (0.6) ST-48 (0.15) ST-45,ST-433, ST-257,ST-658, ST-283 (0.01) others (0.001)
*gln*A21	0.72	none	0.8	ST-52 (0.04) ST-21 (0.03) ST-460 (0.02) ST-354, ST-433 (0.01)
*glt*A5	0.96	ST-49 (1) ST-353 (0.8)	0.2	ST-21 (0.2) ST-353 (0.1) ST-49 (0.09) ST-48 (0.06) ST-257 (0.01) ST-45, ST-22 (0.002)
*gly*A37	0.56	none	0.9	ST-21 (0.03)
pgm_2	0.83	ST-21 (0.87)	0.1	ST-21 (0.7) ST-45 (0.01) ST-48 (0.007) ST-49 (0.004) ST-61, ST-22, ST-353, ST-42, ST-177, ST354 (0.003)
tkt_1	0.96	ST-21 (0.92) ST-48 (0.76)	0.12	ST-21 (0.67) ST-48 (0.15) ST-45 (0.01) ST-257,61,42, 443,573,658 (0.003)
*unc*A5	0.97	ST-21 (0.8) ST-48 (0.9)	0.1	ST-21 (0.6) ST-48 (0.2) ST-61 (0.01) ST-45 (0.008) ST-443 (0.002) ST-177,22,353,42,433 (0.001)
combined *gln*A21 *tkt_1*	1	none	0.96	ST-21 (0.04)

5.3.8 Specific analysis of predictive alleles identified for clonal complex ST-257

The alleles *glyA*62 and *pgm_4* were defined as predictive alleles for identification of this complex. The other alleles identified *asp*A9, *gln*A2, *glt*A4, *tkt_5* *unc*A6 all had high specificities for ST-257 although occurred in many other clonal complexes.

The allele *pgm_4* had a specificity of 0.96, which showed no other high abundance within other clonal complexes, also this allele had a high specificity for this complex from the whole database of 0.7 (Table 5.8). Other low occurrences of this allele were within ST-21, 353, 403 and 45 with low specificity except for ST-21, which had a specificity of 0.7. The allele *gly*A62 showed a high specificity of 0.85 within ST-257. It was initially considered that *gly*A62 could be used independently, however with a small amount of *gly*A62 alleles occurring within clonal complex ST-41, the specificity would have been insufficient. Clonal complex ST-41 is one of the infrequently seen clonal complexes on the database with currently, 19 assigned isolates. However, the good specificity of the allele *gly*A62 of 0.7 for ST-257 in combination with *pgm_4* meant that the combined specificities of 0.98 for ST-257 from the whole database were adequate for assay design. The only occurrence of these alleles together and not in ST-257 was in one unassigned isolate on the database.

Table 5.8 Specificities of alleles for characterisation of clonal complex ST-257 (shaded areas- most applicable alleles and in combination)

Allele	Specificity of allele within ST-257 only	Other high specificity of allele (from table 5.2)	Specificity of allele from whole PubMLST database	
			Assigned to ST-257	Not assigned to ST-257
*asp*A9	0.76	ST-52 (0.7)	0.63	ST-52 (0.2) ST-21 (0.05) ST-45 (0.04) ST-460 (0.01) ST-48 (0.16) ST-354,433,573 (0.0083)
*gln*A2	0.9	ST-42 (0.96) ST-41 (1) ST-362 (1) ST-177 (0.57)	0.3	ST-42 (0.3) ST-41,362 (0.06) ST-177 (0.04) ST-21,353,443 (0.03) ST-206,354,433,49,61 (0.003)
*glt*A4	0.92	ST-573 (0.88) ST-433 (1)	0.58	ST-52 (0.2) ST-21,45 (0.04) ST-460, 48 (0.1) ST-354,433,573 (0.008)
*gly*A62	0.85	ST-41 (1)	0.7	ST-41 (0.18) ST-21, 353, 433, 45 (0.0096)
*pgm*_4	0.96	None	0.7	ST-21(0.7) ST353,403,45 (0.009)
*tkt*_5	0.96	ST-403 (0.78)	0.5	ST-403 (0.4) ST-21,354 (0.006)
*unc*A6	0.93	ST-49 (1) ST-353 (0.9) ST-354 (0.96) ST-658 (0.96) ST-460 (1)	0.5	ST-68 (0.4) ST-21 (0.006) ST-354 (0.006)
combined *gly*A62 *pgm*_4	1	None	0.98	1 unassigned isolate

5.3.9 Summary-identified predictive alleles for each of the target clonal complexes

Predictive alleles for each target clonal complex are summarised (Table 5.9) along with the combined specificities of the alleles from the whole database. When the predictive alleles have been combined for identification of the clonal complexes, very good specificity has been obtained in the range of 0.81 (ST-48) to 1 (ST-61), with the remainder having specificities of 0.96 or higher.

Table 5.9 Predictive alleles for each clonal complex and calculated specificities.

Clonal complex	Potential Alleles identified	Combined specificity from whole MLST database
ST-21	*gln*A1, *tkt*_1	0.98
ST-45	*glt*A10, *tkt*_7	0.99
ST-48	*gln*A4, *unc*A5	0.81
ST-61	*gln*A4, *unc*A17	1
ST-206	*tkt*_1, *gln*A21	0.96
ST-257	*gly*A62, *pgm*_4,	0.98

5.4 General Discussion

By extensive analysis of the data on the PubMLST database it has been possible to characterise a clonal complex by determination of the specific alleles at one or more loci, which can be used as predictive alleles for six target clonal complexes. Initially it was sought to identify one allele, which could be predictive for a complex, although it became apparent throughout these four chapters that the relationships were not specific enough. This was especially the case with the larger clonal complexes (e.g. ST-21) the specificity was insufficient with just one predictive allele, as these were additionally abundant within smaller clonal complexes. With most of the clonal complexes one or more alleles were identified to define a complex, this was especially the case for ST-21, ST-61 and ST-48. The reliance upon two specific alleles would make the strategy more accurate, and avoid the chance of any false positives.

Predictive allele combinations for six clonal complexes have been identified. These six clonal complexes have been previously described to be different with respect to host associations (Colles *et al.*, 2003; Dingle *et al.*, 2002; Manning *et al.*, 2003); therefore it is interesting that alleles of one locus (*gln*A locus) have been used most extensively as predictive alleles, for example in ST-21 (*gln*A1) ST-206 (*gln*A21) and ST-61/48 (*gln*A4). The only locus not incorporated has been the *asp*A locus. These findings are in line with those observed in the study by Dingle *et al* (Dingle *et al.*, 2001), where the alleles of the *asp*A locus were described as having the least number of variable sites and the *gln*A locus having a higher number. The loci, *gly*A, *pgm_* and *tkt_* are described as the alleles with the highest number of variable sites (Dingle *et al.*, 2001), which is in accordance with the fact that the alleles identified at these

loci have been found to be some of the most specific for the clonal complexes, e.g. the high specificity of *pgm*_4 (0.96) for ST-257, and its non-abundance in other clonal complexes.

High specificities were consistent for the combinations of predictive alleles for each clonal complex (Table 5.9). The best specificity (1) was seen for ST-61, which is a result of the highly divergent allele *unc*A17, only found to date in ST-61 assigned isolates and described to have originated in *C.coli* (Colles *et al.*, 2003; Dingle *et al.*, 2001). The specificity for ST-48 although lower than that seen for the other clonal complexes, is adequate for current use (as demonstrated in later chapters). The predictive alleles used for ST-48 also occur within other clonal complexes ST-21 and ST-206. This may be due to the fact that ST-48, ST-21 and ST-206 are described as forming a "complex group" of related genotypes, which are widely distributed (Dingle *et al.*, 2002). Although it has been possible to characterise ST-21 and ST-61 with predictive alleles of high specificities this has not been as straightforward for ST-48. At this point further resolution might have been effectively achieved by complementing the existing predictive alleles with an additional allele. This was not considered because there comes a point where little additional information is attained for the expense and effort involved.

Several other observations concerning the allele distributions were made throughout this section of the investigation. Although the aim was to identify predictive alleles for just six target clonal complexes, it was easy to see from the data analysis carried out here, that it would not be difficult to extend the strategy to include other clonal complexes, especially the smaller more infrequent ones. This was exemplified in the

analysis for ST-257 where the *gly*A62 allele was predictive for ST-257 as well as for the smaller clonal complex ST-41, a similar strategy could be designed for ST-257 and ST-41 as for ST-61 and ST-48 where a common predictive allele was identified for both, and then specific alleles used for differentiating the two complexes.

The methods used to generate this data were all carried out through the use of the search function in the relevant areas of the MLST website. Over the course of this investigation the MLST database was constantly expanding, which had to be taken into account throughout the processes described here. The calculations regarding allele specificities have constantly been updated to account for the expanding database, in spite of this the alleles have remained as specific as originally calculated despite the substantial additions to the database. This provides good evidence that the system is reproducible and robust. If the alleles in which the clonal complexes were highly related to were constantly changing then the frequent amount of extra re-design required for the SNP strategy would vastly outweigh any potential advantages of a strain specific identification assay.

In conclusion it has been possible to identify specific alleles within six target MLST complexes, which could potentially be used for characterisation. A rapid detection method for specific target strains would significantly improve the epidemiological investigation of *C. jejuni.* Therefore the next step of this investigation was to identify the single nucleotide polymorphisms within the predictive alleles (as described in the next chapter).

Chapter 6

An investigation of single nucleotide polymorphisms (SNPs) within locus alleles for strain specific detection and identification of *C.jejuni.*

Chapter 6

An investigation of single nucleotide polymorphisms (SNPs) within MLST locus alleles for strain specific detection and characterisation of *C. jejuni*

6.1 Introduction

The epidemiological paradox of sporadic campylobacter infection is sustained by the absence of routinely used discriminatory sub typing. A novel approach for real time epidemiological characterisation is proposed here. The foremost component to this approach was the identification of stable single nucleotide polymorphisms (SNPs), which identify the allelic profile of six target clonal complexes. In the previous chapter (Chapter 5) predictive alleles were identified, this chapter describes the detection of polymorphic variation by detection of specific SNPs in the allelic sequences which ultimately can be used for assay design.

6.1.1 Single Nucleotide Polymorphisms within the MLST alleles

The accumulation and distribution of allelic SNPs within the *C. jejuni* MLST alleles has been a completely random event, thus the relationship of SNPs to allele number has no ordered structure. MLST allele sequences are given different allele numbers whether they differ at a single nucleotide site or at numerous sites and the number of differences is irrelevant (Dingle *et al.*, 2001). Two consecutively numbered alleles are unlikely to have any relationship in terms of SNPs. The basis behind this is that a single genetic event resulting in a new allele can occur by point mutation (altering a single nucleotide site), by insertion or deletion of foreign DNA or by a recombinational event such as inversion (change in multiple sites).

Most bacterial species have sufficient SNPs within housekeeping genes to provide many alleles per locus, allowing billions of distinct allelic profiles to be distinguished using the seven housekeeping loci. For example: an average of 30 alleles per locus allows approximately 20 billion genotypes to be resolved. For the *C. jejuni* MLST scheme the loci range in size from 402bp (*glt*A locus) to 507bp (*gly*A locus) there have been described to be between 63 (*unc*A locus) and 118 (*pgm_* locus) alleles per locus of which there are variable distributions of polymorphic sites ranging from 63 (15.7%) variable sites in the *glt*A locus to 107 (21.1%) in the *gly*A locus (Dingle *et al.*, 2001). This variability within the alleles was harnessed in the approach described herein. Informative SNPs, in the predictive alleles (described in the previous chapter) were identified to ultimately determine the allelic profile of the major epidemiological lineages (clonal complexes) ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257 and provide a basis for the design of allelic discrimination assays.

6.2 Materials and Methods

6.2.1 Downloading of MLST Alleles

Using the Allelic profile/ST Database (http://pubmlst.org/campylobacter) and "Download Alleles" function, all alleles at each locus (excluding *asp*A) were downloaded in FASTA format into Bioedit Sequence Alignment Editor version 4.0.9 (copyright Tom Hall, Dept of Microbiology, North Carolina State University).

6.2.2 Allele Alignments

The predictive allele was placed at the top to be used as the consensus, and all sequences at the locus were aligned using the ClustalW multiple alignment feature in Bioedit. All allele sequences at this locus were plotted to the consensus with a dot when the base was identical, or shown as the base designation where different. This enabled the informative SNPs within the predictive allele to be easily identified.

6.2.3 Nucleotide Comparisons and Identification of Polymorphisms

All areas of allelic variation were identified by eye and with the use of information on the MLST website describing the location of the polymorphic sites within alleles. SNPs unique to the predictive allele were identified. For each predictive allele up to four SNPs were identified, where possible on the following criteria;

(i) Exclusivities for the predictive allele

(ii) Within a central region of the allele sequence

(iii) Within an area of even base ratios

(iv) Where flanking regions were relatively conserved

6.3 Results and Discussion

6.3.1 Identification of informative SNPs within the MLST predictive alleles

"Informative SNPs" were the polymorphisms present within the predictive alleles, which were most exclusive for the identified alleles and occurred in very few other alleles. This enabled the characterisation of a predictive allele based upon its well-defined polymorphisms, which were a distinct feature of the allele sequence. For most of the predictive alleles identified, more than one SNP per allele was required, therefore by utilising a combination of SNPs within different predictive alleles, specificity for a clonal complex could be obtained.

Six separate sequence alignments *gln*A, *glt*A, *gly*A, *pgm_*, *tkt_* and *unc*A were constructed to identify the informative SNPs (no *asp*A predictive alleles had been identified in the previous chapter, therefore *asp*A was excluded). These were adjusted where required to accommodate different consensus sequences. For example a different *gln*A predictive allele was used in the identification of four clonal complexes (ST-21, ST-61/48 and ST-206). The number of SNPs identified per allele was very dependent on the nature of the allele sequence. In the alleles, which showed a high degree of variability, for example, *unc*A17, a large number of exclusive informative SNPs were present, therefore ones chosen were towards the centre of the allelic sequence and within an area of balanced base ratios. The allele *unc*A17 as described in the previous chapter showed a high degree of variability, which may have originated in a different *Campylobacter* species (Dingle *et al.*, 2001). In less variable alleles, for example *gln*A1, fewer choices in informative SNPs were available therefore the most exclusive were used. For example the informative SNP within the allele *gly*A62 (for ST-257) was towards the very end of the allele

sequence, which could be problematical for detection by PCR based methodologies, however this was the most informative SNP available.

6.3.2 Identification of Informative SNPs for the ST-21 Clonal Complex

For identification of clonal complex ST-21 the predictive alleles *gln*A1 and *tkt*_1 had been identified in the previous chapter. The *gln*A locus had a length of 477bp and the alignment included 103 sequences, with *gln*A1 as the consensus the overall alignment was relatively conserved with very few SNPs present amongst the lower allele numbers but with the occurrence of more SNPs within the later numbers. Extensive analyses of the alignment for the *gln*A locus, with the allele *gln*A1 used as the consensus, enabled the identity of two SNPs. A bialleleic SNP A→G was identified at 108bp, with the base A within the *gln*A1 and G in 90% of the other alleles (Appendix 2.1). This SNP was not exclusive for *glnA*1 and a further SNP was identified to determine the allele. The second SNP identified was a C at 267bp (Appendix 2.2). This was a triallelic SNP, which also occurred as an A, or T within other alleles. By combining the two SNPs the allele *gln*A1 could be identified from the majority of the other alleles (Table 6.1).

Comparison of *gln*A alleles showed that these two SNPs were also present in *gln*A alleles 14, 20, 56, 58 and 80 (Table 6.1). These were investigated further to determine how frequently they occurred and within which complexes. The low occurrence of these alleles was ignored as they all occurred within ST-21 apart from *glnA*20, which had 2 occurrences within ST-206, however this would be overcome by the use of *tkt*_1 for identification of this complex.

The areas in which these two SNPs were present differed. The SNP within the region of 108bp occurred in a largely conserved region apart from a SNP C→T, which occurred at 90bp in a number of alleles. Also another SNP C→T in this area occurred at 111-113 bp. Conversely the regions around the second SNP at 267bp were less conserved especially in the later alleles (*gln*A37 and above) where large amounts of sequence diversity were present in the areas surrounding this SNP.

The *tkt_* locus had a length of 459bp and the alignment included 111 sequences where greater variation between the alleles was evident. By alignment of the *tkt_* alleles with *tkt*_1 as the consensus, a triallelic SNP T→C/G was identified at 330bp, which occurred as a T in *tkt*_1 and a C or G in other alleles (Appendix 2.3). This SNP occurred in areas with much sequence variation and the occurrence of many other SNPs. As well as in *tkt*_1 this SNP was also present within the alleles, *tkt*_4, 27, 53, 78, 90 and 100, which occurred in the ST-21 clonal complex. However this SNP occurred in other alleles outside the ST-21 clonal complex; *tkt*_39 in ST-206, *tkt*_62 in ST-48, *tkt*_92 in ST-48 and *tkt*_101 in ST-45. These were all low occurrences except for the allele *tkt*_62, which showed 24 occurrences within the clonal complex ST-48. To determine if this was going to affect the specificity a search was made on the database to determine if there was any possibility of this allele occurring in combination with the allele *gln*A1. This was determined not to be an issue and the alleles *gln*A1 and *tkt*_62 had to date, never been assigned into a complex together (Table 6.1). The presence of the SNP in other alleles not described was overlooked for the purpose of this assay design, as other sequence variation in alternative positions in close proximity to the SNP was present. This was the case for the allele *tkt*_35, where the SNP was present, however there was a large amount of

sequence variation surrounding this SNP (320-325bp), this is explained in greater detail in the next chapter.

Table 6.1 Identified SNPs within each predictive allele for the ST-21 clonal complex.

Allele	Position of SNP (bp) on allele and base change	Other alleles with these SNPs and the clonal complex (no. of entries on database)*
*glnA*1	108 (A→G) 267 (C→T/A)	*glnA*14: (none in database) *glnA*20: ST-21 (5), ST-206 (2) *glnA*56: ST-21 (1) *glnA*58: ST-21 (1) *glnA*80: ST-21 (1)
*tkt*_1	330 (T→C/G)	*tkt*_4: ST-21 (1) *tkt*_27: ST-21 (1) *tkt*_39: ST-206 (1) *tkt*_53: ST-21 (1) *tkt*_62: ST-48 (24) *tkt*_78: ST-21 (1) *tkt*_90: ST-21 (2) *tkt*_92: ST-48 (2) *tkt*_100: ST-21 (1) *tkt*_101: ST-45 (1)

*** data calculated 05/12/03**

6.3.3 Identification of SNPs for the ST-45 clonal complex

For identification of clonal complex ST-45 the predictive alleles *glt*A10 and *tkt*_7 had been identified in the previous chapter. The *glt*A locus was the shortest allele with a length of 402bp and the alignment included 79 allele sequences. This was generally conserved except for variation around 205-215bp in the higher numbered alleles. Within the allele *glt*A10 two SNPs were identified which could potentially distinguish this allele. These were a triallelic SNP C→G/A at 201bp and a further biallelic C→T at 225bp (Appendix 2.4). By combining these two SNPs the allele *glt*A10 could be identified as they only occurred in the other alleles *glt*A14, 40, 41, 51, 70 and 72. As in the case for ST-21 these alleles were investigated further to determine if the clonal complex specificity was going to be reduced by the possibility of also detecting these extra alleles in other clonal complexes. The majority of these alleles were present with low occurrence and were assigned into the ST-45 clonal complex. However *glt*A40 had 40 occurrences within ST-283, which would obviously affect the specificity of the assay (Table 6.2). Consequently, as in the case of ST-21, further SNPs within the second predictive allele *tkt*_7 were used to improve the specificity.

The SNPs identified within *tkt*_7 (Appendix 2.5) were the triallelic A→ C/T at 138bp and biallelic C→T at 141bp. The distribution of these SNPs was constant throughout the other alleles with most alleles showing both base changes. Exceptions included *tkt*_86, *tkt*_104, and *tkt*_106-108 where only one base change was present at the position 141bp. As well as identifying *tkt*_7 these SNPs were present in alleles *tkt*_10, 14, 20, 28, 30, 33, 51, 73, 76 and 110. The only allele, which was likely to affect the specificity, was the allele *tkt*_28, which showed 49 occurrences within the

clonal complex ST-508. To determine the specificity of this combined with the allele *glt*A10 searches were made on the database. Only one entry on the database contained the alleles *tkt*_28 and *glt*A10, which was assigned into the ST-45 clonal complex.

Table 6.2 Identified SNPs within each predictive allele for the ST-45 clonal complex.

Allele	Position of SNP (bp) on allele and base	Other alleles with these SNPs and Clonal complex (entries on database)*
***gltA*10**	201 (C→G/A) 225 (C→T)	*gltA*14: ST-658 (1), ST-45 (5), ST-42 (1) *gltA*40: ST-283 (40), ST-45 (7) *gltA*41:ST-45 (4) *gltA*51: ST-45 (2) *glt*A70:ST-45 (2) *glt*A72:unassigned (2)
tkt_7	138 (A→C/T) 141 (C→T)	*tkt*_10 :ST-21 (1), unassigned (1) *tkt*_14 :ST-45 (1) *tkt*_20 :ST-45 (1), ST-21 (1) *tkt*_28 :ST-508 (49), ST-45 (1) *tkt*_30 :ST-206 (1), ST-45 (1) *tkt*_33 :none *tkt*_51 :ST-283 (22), ST-45 (10), ST-48 (10), ST-206 (1) *tkt*_73 :unasigned (1) *tkt*_76 :unassigned (10) *tkt*_110: none

*** data calculated 05/12/03**

6.3.4 Identification of SNPs for the ST-61 and ST-48 clonal complexes

A combined strategy was used for the identification of the two clonal complexes ST-61 and ST-48 due to the occurrence of the highly abundant predictive allele *gln*A4 in both (as described in the previous chapter). The two complementary alleles *unc*A5 and *unc*A17 were identified as predictive alleles for clonal complexes ST-48 and ST-61 respectively.

For the allele *gln*A4, used to initially determine both clonal complexes ST-61 and ST-48, two SNPs were identified. A biallelic SNP C→T at 18bp was identified (Appendix 2.6). This SNP was towards the very start of the allele sequence, which could have been difficult for detection (as described later), however this was one of the most informative SNPs present. A second biallelic SNP G→A was identified at 202bp (Appendix 2.7) towards the centre of the sequence. These two SNPs were also present in other alleles including *gln*A32, *gln*A51 *gln*A63 and *gln*A102 (Table 6.3). The alleles predominantly occurred within the clonal complexes ST-61 and ST-48, except for one occurrence of the allele *gln*A32 in ST-52 and one occurrence of *gln*A63 in ST-21, therefore these were unlikely to decrease the specificity when included in conjunction with other alleles.

The allele *unc*A5 had been identified as a predictive allele for ST-48, which was required to achieve adequate specificity for this clonal complex. The *unc*A locus had a length of 489bp and the alignment for this locus included 63 sequences. Two SNPs at 186bp A→T and 189bp T→C/A were identified for the *unc*A5 allele (Appendix 2.8), these SNPs were not exclusive to *unc*A5 and also occurred in other alleles within clonal complex ST-48 including *unc*A25 (1 isolate), and outside of clonal

complex ST-48, *unc*A10 (ST-21, 1 isolate), *unc*A20 (ST-61, 1 isolate), *unc*A25 (ST-61, 2 isolates and ST-508, 1 isolate) *unc*A59 (unassigned, 1 isolate) and *unc*A63 (unassigned, 2 isolates) however these were of low occurrence. Additionally, these SNPs also occurred in the alleles *unc*A11 and *unc*A30 however no isolates with these alleles had been described on the database. In conjunction with the use of the allele *gln*A4, the required specificity could be achieved.

To complete the identification of ST-61 clonal complex the predictive allele *unc*A17 had been identified. The alignment of the *unc*A17 allele against the rest of the *unc*A alleles looked very different from other comparisons (Appendix 2.9). This was due to the highly diverse nature of allele *unc*A17 (Dingle *et al.*, 2001) and the high number of variable sites. Numerous SNPs could have been used, however a region containing three SNPs was identified between 330bp and 339bp. This region was chosen on the basis that the flanking regions were relatively conserved enabling the design of assays (as described in the next chapter) easier. Additionally it was located in a central region of the allele and the ratio of bases was even. Other alleles were identified by detection of these SNPs but these mainly resided in clonal complex ST-61 or were unassigned. The only allele that occurred in another clonal complex (ST-283) was *unc*A38, which had one entry on the database.

Table 6.3 Identified SNPs within each predictive allele for the ST-48 and ST-61 Clonal complexes.

Allele	Position of SNP (bp) on allele and base change	Other alleles with these SNPs and Clonal Complex (entries on database)*
glnA4	18 (C→T) 202 (G→A)	*gln*A32: ST-48 (1), ST-61 (1), ST-52 (1) unassigned (1) *gln*A51: ST-61 (1) *gln*A63:ST-21 (1) *gln*A102: ST-48 (1)
***unc*A5**	186 (A→T) 189 (T→C/A)	*unc*A10: ST-21 (1) *unc*A11: none *unc*A20: ST-61 (1) *unc*A25: ST-48 (1), ST-61 (2), ST-508 (1) *unc*A30: none *unc*A59: unassigned (1) *unc*A63: unassigned (2)
***unc*A17**	333 (G→A) 336 (C→T) 339 (T→A)	*unc*A28: unassigned (1), ST-61 (1) *unc*A36: unassigned (1) *unc*A37: none *unc*A38: ST-61 (2), ST-283 (1) *unc*A41: none *unc*A42: ST-61 (3)

*** data calculated 05/12/03**

6.3.5 Identification of SNPs for the ST-206 clonal complex

For identification of clonal complex ST-206 the predictive alleles *gln*A21 and *tkt*_1 had been identified in the previous chapter. The SNPs used for the identification of *tkt*_1 (Section 6.3.2) were also used for identification of this clonal complex in conjunction with two further bialleleic SNPs on the *gln*A21 allele. These were a C→T at 18bp and an A→G/T SNP at 33bp (Table 6.4). The overall consensus alignment for *gln*A21 showed a wide distribution of SNPs within otherwise relatively conserved regions. The amount of variation increased among the higher number alleles (Appendix 2.10).

These two SNPs were almost exclusive for *gln*A21; only *gln*A64 and *gln*A84 also contained these SNPs. These alleles were investigated to determine their occurrence (Table 6.4), where *gln*A64 had one occurrence on the database assigned into ST-206 and *gln*A84 had no entries.

Table 6.4 Identified SNPs within each predictive allele for the ST-206 clonal complex.

Allele	Position of SNP on allele and base	Other alleles with these SNPs and the ST-complex they occur (entries on database)*
*tkt*_1	330 (T→C/G)	*tkt*_4: ST-21 (1) *tkt*_27: ST-21 (1) *tkt*_39: ST-206 (1) *tkt*_53: ST-21 (1) *tkt*_62: ST-48 (24) *tkt*_78: ST-21 (1) *tkt*_90: ST-21 (2) *tkt*_92: ST-48 (2) *tkt*_100: ST-21 (1) *tkt*_101: ST-45 (1)
*gln*A21	18 (C→T) 33 (A→G/T)	*gln*A64: ST-206 (1) *gln*A84: no entries

*** data calculated 05/12/03**

6.3.6 Identification of SNPs for the ST-257 clonal complex

For identification of clonal complex ST-257 the predictive alleles *gly*A62 and *pgm*_4 had been identified in the previous chapter. The *gly*A locus had a length of 507bp and the alignment included 111 sequences. A biallelic SNP C→T was identified at 483bp, which enabled the allele *gly*A62 to be distinguished. This SNP within the allele *gly*A62 (for ST-257) was towards the very end of the allele sequence, which could be problematical for detection by PCR based methodologies, however this was the most informative SNP available. Other alleles with this SNP were *gly*A63, *gly*A87, *gly*A99 and *gly*A109. *Gly*A63 was assigned to ST-257, and the remaining alleles only occurred in unassigned entries (Appendix 2.11) (Table 6.5).

The *pgm*_ locus had a length of 498bp and the alignment included 118 sequences. The allele *pgm*_4 contained a polymorphic site between 162bp and 171bp, which was used for identification (Appendix 2.12). Within this region there were up to four SNPs, which could be used to identify the allele, however these were present in different combinations in the other alleles. The majority of the other alleles contained all four SNPs, however for example *pgm*_16 only contained the one base change at 165bp. The number of alleles in which these SNPs also occurred was numerous, however by combining the allele *pgm*_4 with *gly*A62 then adequate specificity was achieved. This was due to the very low occurrence of allele *gly*A62 outside ST-257, but the high prevalence of the allele *pgm*_4 within ST-257, despite the high occurrence within other complexes. Identification of this allele highlighted the problem that arises when stable polymorphisms were not sufficiently frequent to enable easy detection of the allele by one or two SNPs. Further resolution may have been effectively achieved by a complementary SNP, however little additional

information would have been achieved for the effort required, as specificity was adequate for current use, complementary SNPs were not considered.

Table 6.5 Identified SNPs within each predictive allele for the ST-257 clonal complex.

Allele	Position of SNP (bp) on allele and base change	Other alleles with these SNPs and Clonal Complex (entries on database)*
***glyA*62**	483 (C→T)	*glyA*63: ST-257 (1) *glyA*87: unassigned (2) *gly*A99: unassigned (1) *gly*A109: unassigned (1)
pgm_4	162 (G→A/T) 165 (C→A/T/G) 168 (A→T/G) 171 (G→A)	*pgm_6:* ST-61 (13), ST-48 (22), ST-403 (15), ST-21 (2), ST-177 (1) ST-443 (1). *pgm_11*: ST-49 (27),ST-354 (26), ST-353 (16), ST-362 (16), ST-48 (14), ST-52 (6), ST-21,22, 460, 573, 658 (1) *pgm_13*: ST-353,354,658 (1) *pgm_14*: ST-354, 21 (1), *pgm_17*: (26) *pgm_22*:ST-52 (23), ST-21 (3), ST-257 (1) *pgm_23*: ST-443 (23), ST-206 (3),ST-48,52 (1) *pgm_28*: none *pgm_29*: ST-61, 42 (1) *pgm_44* ,55: none *pgm_56* :unassigned (1) *pgm_66* :ST-433 (1) *pgm_67*:ST-353 (2) *pgm_74*:ST-48 (3) *pgm_77*,79:unassigned (1) *pgm_82* ST-354 (1) *pgm_83*: ST-21 (1), unassigned (1) *pgm_86*: ST-52 (5), unasigned (1) *pgm_89* :ST_460 (5), unassigned (1) *pgm_90*:ST-362 (1), unassigned (1) *pgm_95*:ST-52 (2), ST-48 (1)

*** data calculated 05/12/03**

6.3.7 General Discussion

Single Nucleotide Polymorphisms (SNPs) providing rapid multi-allelic profiles, diagnostic of important lineages and amenable to real time sequence analysis have been identified. These SNPs detected with the appropriate instrumentation will ultimately enable direct detection of specific *C. jejuni* strains by clonal complex. There are many new platforms available for the detection of short sequences and SNPs, as described in the next chapter, which would provide accurate and rapid results.

The strategy used, confirmed that informative SNPs were present within the predictive alleles, which could be used for characterisation. The alignments in Bioedit were accurate and allowed the straightforward identification of SNPs unique to each predictive allele. The six target clonal complexes have been identified by the use of ten predictive alleles encompassing a total of 21 SNPs (Table 6.6). This strategy for the six clonal complexes was made easier by the sharing of predictive alleles (and associated SNPs) for some clonal complexes including the allele *gln*A4 for ST-48 and ST-61 and the allele *tkt*_1 for ST-21 and ST-206.

All of the polymorphic sites seen within the alleles investigated were distributed randomly along the gene sequences, consistent with evolution by point mutation (Viscidi & Demma 2003). The presence and positions of SNPs resulting from the extensive genetic exchange as well as intragenomic alterations which occur within *C. jejuni* populations *in vivo* (de Boer *et al.*, 2002) are the key components for the diversity seen within *C. jejuni* isolates analysed by MLST. Interestingly within the alignments constructed for the purposes of this chapter, many "hotspots" with

widespread sequence divergence, were identifiable in the higher number alleles. This may reflect the expansion of the database from isolates associated with human disease to other sources. It would be interesting to examine these areas in greater detail to determine if there was any relationship between presence of SNPs and source.

Sufficient genetic diversity was present in the majority of the alleles across all the loci to make it possible to identify polymorphic sites, which could be used to distinguish a single allele. Detection by the means of a single SNP was not feasible for some of the key alleles, so a strategy using a combination of SNPs was adopted. This approach ensured that strains would not be incorrectly assigned to clonal complexes on the basis of the detection of one SNP. A multi SNP approach increased the certainty that the clonal complex identified was correct and would reduce the number of false positives. Additionally the chance of identifying novel allele sequences may be facilitated by the detection of multiple SNPs on the same allele and the use of more than one allele. The more points on the allele detectable then the risk of confusion by the chance association of SNPs was decreased.

It has been demonstrated that for the six major clonal complexes it has been possible to identify allelic sequences, which can be used for identification of a clonal complex. This reflects the sequence of the alleles and their groupings into lineages. Alleles are assigned arbitrary numbers in order of identification, therefore the alleles identified by detection of a particular SNP will not be related or identified by consecutive numbers. The emphasis when choosing the SNPs, was on the detection of the particular allele which occurred in as few as possible of the more common

alleles. In many cases the occurrence of the SNP outside the clonal complex was in alleles with one occurrence on the database and usually unassigned into a clonal complex. For example, for the allele *gly*A62 the other alleles containing this SNP were *gly*A63 (ST-257, 1 isolate), *gly*A87 (unassigned, 2 isolates), *gly*A99 (unassigned, 1 isolate), *gly*A109 (unassigned, 1 isolate).

The variation seen with respect to the number of SNPs present within the alleles was low considering that *C. jejuni* is a weakly clonal species. This can be seen by comparison with other organisms for which MLST schemes have been established, for example when compared to *N. menigitidis*. Reasons for such a small amount of sequence diversity in the *C. jejuni* genome, could be attributed to the fact that *C. jejuni* is a young species and therefore has not had sufficient time to accumulate sequence diversity (Suerbaum *et al.*, 2001). The almost complete lack of sequence diversity in *Y. pestis* has been attributed to the fact that this species has emerged as a new clone of *Y. pseudotuberculosis.* Another possible reason for the lack of sequence variation is that the population of *C. jejuni* has recently undergone expansion, as a result of changes in its ecology. This may have resulted from changes in human food preparation especially in relation to slaughtering practices. This expansion of clones to fit into the changes of food husbandry could explain the limited number of alleles and low frequency of synonymous nucleotide polymorphisms (Suerbaum *et al.*, 2001).

This approach for the identification of SNPs within alleles represents a significant generic method, which could be applied for delivery of real time epidemiology for other pathogens. The approach for identification of SNPs within defined alleles could

be applied for sequence-based studies other than those utilising MLST. Similar approaches have been described for organisms where rapid typing or detection methods have been implemented based on the presence of SNPs. For example in *L. monocytogenes,* SNPs have been identified within the *sig*B gene to differentiate different lineages (Moorhead *et al.*, 2003) or to discriminate within the *Listeria* genus (Koo & Jaykus, 2002). Similarily, SNPs have been mapped within the type III secretion toxins of *Pseudomonas aeruginosa* for the development of a multiplex PCR system (Ajayi *et al.*, 2003). A SNP based strategy was also developed for examining the distribution of biomedically relevant traits such as virulence and transmissibility of the *M.tuberculosis* complex organisms (Gutacker *et al.*, 2002). In addition SNPs have been used for rapid molecular identification and subtyping of *H.pylori* by determination of SNPs within the 16S rDNA variable V1 and V regions (Monstein *et al.*, 2001).

The allelic polymorphisms from MLST gene sequences identified here allow for the subsequent development of allelic discrimination assays for potential rapid strain profiling. The evolution of the data presented in this chapter into assays for the detection of sequence polymorphisms applicable for rapid strain detection and characterisation, by two different methods is described in the next two chapters.

Table 6.6 Summary of Alleles and associated SNPs used for identification for the six target clonal complexes

Clonal Complex	Predictive Alleles	SNPs (bp)
ST-21	*gln*A1	108 (A→G) 267 (C→T/A)
	*tkt*_1[1]	330 (T→C/G)
ST-45	*glt*A10	201 (C→G/A) 225 (C→T)
	*tkt*_7	138 (A→C/T) 141 (C→T)
ST-48	*glnA4*[2]	18 (C→T) 202 (G→A)
	*unc*A5	186 (A→T) 189 (T→C/A)
ST-61	*glnA4*[2]	18 (C→T) 202 (G→A)
	*unc*A17	333 (G→A) 336 (C→T) 339 (T→A)
ST-206	*tkt*_1[1]	330 (T→C/G)
	*gln*A21	18 (C→T) 33 (A→G/T)
ST-257	*glyA*62	483 (C→T)
	pgm_4	162 (G→A/T) 165 (C→A/T/G) 168 (A→T/G) 171 (G→A)
Totals for 6 clonal complexes	10 predictive alleles	21 SNPs

[1] ***tkt*_1 allele and SNPs used for both ST-21 and ST-206**
[2] ***gln*A4 allele and SNPs used for both ST-48 and ST-61**

Chapter 7

Design of Assays for strain specific detection and characterisation based on the Roche Lightcycler.

Chapter 7

Design of Assays for strain specific detection and characterisation based on the Roche Lightcycler

7.1 Introduction

Various technologies have been developed for rapid sequence analysis involving SNPs, which could provide the rapid multi-allelic sequence profiles diagnostic of important *C. jejuni* lineages. The next step in the process was the design and testing of MLST SNP assays based on the Roche Lightcycler for rapid strain characterisation.

Many methods for identification of SNPs have been based on restriction fragment length polymorphism techniques (Moorhead *et al.*, 2003) or full sequencing (Read *et al.*, 2002) which are both lengthy and time-consuming processes. More recently newer and faster technologies, with specialised platforms for SNP detection have become available. Existing systems, which can be applied for SNP detection include mass spectroscopy (Walters *et al.*, 2001), primer extension assays (Rudi & Holck, 2003; Ye *et al.*, 2001), pyrosequencing (Monstein *et al.*, 2001; Unnerstad *et al.*, 2001) and real time PCR approaches using the Taqman or Lightcycler (Jurevic *et al.*, 2003; Stermann *et al.*, 2003). Alternative approaches are the newer specialist type of system designed purely for high throughput detection of SNPs such as the Roche Lighttyper (www.lightyper.com).

SNP Detection on the Lightcycler

Various platforms for detection of SNP's are available which offer rapid and accurate determination of SNPs. The Lightcycler was chosen as a suitable platform for the implementation of assays, due to the instrument's ability and reported accuracy of detecting SNPs through the melt curve program (Lohmann *et al.*, 2000). The machines are also strategically placed in both research and diagnostic laboratories and also throughout many laboratories worldwide making it possible to roll out rapid and comparable genetic characterisation in multiple locations.

The melting curve format provides a powerful means for rapid detection of SNPs, based on melting peak analyses, which monitor the temperature dependent hybridisation of probes to single stranded DNA. The temperature required to separate the hybridisation probes from target DNA is not only dependent on length and G+C content but also on the degree of homology (Reiser *et al.*, 1999). Mismatches between the probe and target sequence cause destabilisation resulting in a lower melting temperature than what would be required to separate the probe and target if the DNA were perfectly matched (Anon., 2003).

The Lightcycler exploits these melting temperature differences in order to detect the presence/absence of a SNP. Two probes are used; the sensor probe covers the region of potential mutation and has a lower T_m than the adjacent probe, the anchor probe. This ensures that the fluorescent signal generated during the melting curve is determined only by the sensor probe (Reiser *et al.*, 1999). Probes binding with target DNA containing a mismatch will melt at a lower temperature than those with perfectly matched DNA, enabling subsequent identification of a SNP, by the higher

T_m. The Lightcyler system also allows simultaneous identification of up to two SNPs in the same capillary. Dual colour detection can be used with two different hybridisation "sensor" probes with different fluorescent labels; one sensor probe is labelled with LC Red 640 and the other labelled with LC Red 705 (Anon., 2003).

Results from the previous two chapters were used to design and implement assays based on the detection of informative SNPs within predictive alleles of six MLST clonal complexes. It was sought to design a series of Lightcycler assays in either a uniplex or duplex format depending on the numbers and types of SNPs to be identified, to rapidly screen isolates for their clonal complex based on the presence of specific SNPs. The original MLST reference isolates described by Wareing *et al* (Wareing *et al.*, 2003), were used for validation together with a panel of MLST characterised isolates from animal, human and environmental sources.

7.2 Materials and Methods

7.2.1 Isolates Selected

Set 1, *C. jejuni* MLST Reference isolates

A set of reference isolates for the main clonal complexes from the *C. jejuni* MLST scheme (Wareing *et al.*, 2003) are available through the National Collection of Type Cultures (NCTC London, UK). These were chosen using information from the population structure gained from MLST and represented the founder genotypes of 13 clonal complexes, which were most commonly identified among a diverse collection of isolates. These isolates were originally obtained from the UK and Netherlands from a range of sources including human disease, poultry, cattle, sheep, wild birds, and sand from bathing beaches. These isolates were used for the initial implementation and evaluation of the assays. Two additional strains were included in this study, these being the founder strains for clonal complexes ST- 403 and ST-353.

Set 2, Panel of diverse human and environmental samples

Two hundred and twenty one isolates from the Campylobacter Reference Unit Culture Collection were chosen as being representative of isolates from different clinical and environmental sources. These were independent from isolates reported in any other *C. jejuni* MLST publication. They included strains from a UK wide retail poultry survey (n=88)[1], from human enteritis, referred as part of the Campylobacter Sentinel Surveilance Scheme (n=90) (Gillespie *et al.*, 2001), from animals (dogs, birds, cattle) (n=22), and from animal products (n=21).

[1] (www.foodstandard.gov.uk/multimedia/ pdfs/campsalmsurvey.pdf)

7.2.2 Full MLST of the reference isolates and isolates (n=221) from various sources of infection.

The full MLST procedure was carried out in the Department of Zoology, University of Oxford using an ABI 3700 Sequencer, or in the Laboratory of Enteric Pathogens Health Protection Agency, London, with the adapted method for the Beckman CEQ8000 Capillary Sequencer. Both methods are described in detail in chapter 2.

Data analysis was carried out using two different software applications. In Oxford the STARS (Sequence Typing Analysis and Retrieval System) was used which allows the automatic assembly of short contigs into the same database and automatic designation of the alleles. In London, manual assembly of contigs and editing was performed using Bioedit, with manual assignment of allele types using the MLST website. With both systems STs and clonal complexes were assigned by interrogation of the database as described in section 2.

(I wish to acknowledge the help of Martin Maiden, Frances Colles, Roisin Ure (University of Oxford) and Barry Curran (University of Warwick) in this section for their help in generation of the full MLST profiles.)

7.2.3 Design of Roche Lightcycler assays

Materials

- Lightcyler probe Design Software (Version 1.0 Roche diagnostics)
- Sequence of MLST predictive alleles and positions of SNPs (identified in chapters 5 and 6).

Method

Primer and probe combinations were designed using the Lightcycler probe design software or manually (by eye) and with the use of a T_m calculator (T_m utility Version 1.5, iTech). The criteria specified within the software and by the Lightcycler guidelines were used. Assays were designed in either a single or duplex format dependent on the number of SNPs to be detected. For each SNP two primers, one anchor probe and one sensor probe were designed. The sensor probe was placed over the area of the SNPs and anchor probe and primers were placed in areas with conserved sequence, to enable amplification from all allele sequences, regardless of the allele type. In the duplex reactions where the two SNPs were sufficiently close together, only one set of primers was required. All primer and probe sequences were checked in BLASTN (http://www.ncbi.nlm.nih.gov/BLAST/) to confirm specificity for *C. jejuni*.

7.2.4 Lightcycler amplification and melting peak analysis

Materials

- Lightcyler Instrument (1.2 Instrument with version 3.5 Lightcycler software)
- Lightcycler Faststart hybridisation probes mastermix (Roche)
- 0.3μM primers, forward and reverse (MWG Biotech)
- 0.1μM sensor probes labelled with 5' LCred 640 or 705 and phosphorylated at the 3'end (Metabion, Planegg-Martinsried, Germany).
- 0.1μM anchor probes labelled with 3' fluorescein (Metabion)
- 25mM $MgCl_2$ stock.
- Glass capillaries and with plastic stoppers (Roche)
- LC Centrifuge (Roche)

Method

Lightcycler reactions were initially applied to isolated DNA and thereafter to MLST locus amplified first round PCR products (as described in section 2). Lightcycler reaction conditions were originally set up as described by the manufacturer and were subsequently optimised to determine the best conditions for reaction efficiency with the different primer and probe combinations.

The Roche Lightcycler was used for all reactions. Lightcycler reaction mastermix contained, per 10µl reaction, 0.3µM primers, forward and reverse, 0.1µM sensor probes labelled with 5' LCred 640 or 705 and phosphorylated at the 3'end, 0.1µM anchor probes labelled with 3' fluorescein, 2.5µl Lightcycler Faststart hybridisation probes mastermix, 2-4 mM $MgCl_2$ and 1-2µl 1st round PCR product. All reagents were added into glass capillaries and sealed with plastic stoppers. $MgCl_2$ was titrated, to find optimal concentrations for each reaction.

Optimal conditions were determined to be an initial denaturation step at 95°C for 10 minutes, followed by 40 cycles of 95°C for 10 seconds, 63°C for 10 seconds and 72°C for 10 seconds (ramp rate 20:00). The melting steps were adjusted for optimal melting peaks depending on the assay and consisted of denaturation of 95°C for 1 second, followed by 45-53°C for 1 minute, and slowed extension with ramp rate of 0.1°C/second to 85°C, with continuous data acquisition. Fluorescence was detected by the instrument's fluorimeter and analysed using Lightcycler software. Melting peaks were derived automatically by plotting the first negative derivative (-dF/dT) of the fluorescence against the temperature, which revealed the differences between the melting temperatures of the amplified products (Reiser *et al.*, 1999).

7.3 Results and Discussion

7.3.1 General assay implementation

In total, six Lightcycler assays were designed; these comprised a duplex and uniplex assay for identification of clonal complex ST-21, a uniplex assay for ST-257 and duplex assays for ST-45, ST-61 and ST-48. The whole complement of predictive alleles and associated SNPs identified in chapters 5 and 6 was not used, and a strategy based on seven predictive alleles was initially devised and tested. For every assay, $MgCl_2$ was titrated from 1-5mM to find the optimal concentration at which the majority of the assays were optimal at 2-4mM $MgCl_2$ (Table 7.1). In addition, each assay was optimised with the use of PCR products and by variation in the annealing time and temperature, during the amplification step and the melting process (Final reaction conditions and primer and probe sequences are shown in tables 7.2 and 7.3).

Table 7.1 $MgCl_2$ final concentration used for each assay

Clonal Complex	Assay	$MgCl_2$ concentration (mM)
ST-21	*gln*A1	4
	*tkt*_1	4
ST-45	*glt*A10	4
ST-48&61	*gln*A4	4
	*unc*A5 *unc*A17 (duplex)	2
ST-257	*gly*A62	3

In the early stages of assay development, 10ng of MagNApure *C. jejuni* extracted DNA was added into 10µl Lightcycler reactions. However it was determined despite numerous attempts at varying the concentrations of DNA, primers, probes,

mastermix and reaction conditions, that it was not going to be possible to successfully implement these Lightcycler SNP assays on extracted genomic DNA. Attempts were made to resolve extraction problems by heating of the DNA to 95°C for ten minutes in order to release the supercoiling. DNA was purified by the use of spin columns (Amersham) and serial dilution of the samples. Despite testing a wide range of options, assays using extracted genomic DNA were unsuccessful. Subsequently the assays were applied to first round PCR products. These were more effective and made the development of specific assays by use of the nested approach possible.

Table 7.2 Reaction conditions for each Lightcycler assay. Red value indicating variation in annealing temperature.

Clonal Complex	Predicitive allele (SNPs)	De naturation			Amplification			Melting Curves			Detection Channel
		Temp (°C)	Time (min:sec)	Temp trans. (°C/sec)	Temp (°C)	Time (min:sec)	Temp trans. (°C/sec)	Temp (°C)	Time (min:sec)	Temp trans. (°C/sec)	
ST-21	*gln*A1 (108,267)	95	10:00	20:00	95 61 72	15 5 10	20:00 20:00 10:00	95 53 72	1 1:00 0	10:00 20:00 0	F2, F3
	*tkt*_1 (330)				95 61 72	15 5 10	20:00 20:00 20:00	95 51 85	1 1:00 0	10:00 20:00 0.1	F3
ST-45	*glt*A10 (201,225)				95 63 72	10 10 10	20:00 20:00 20:00	95 50 85	1 1:00 0	10:00 20:00 0.1	F2, F3
ST-257	*gly*A62 (483)				95 63 72	10 10 10	20:00 20:00 20:00	95 50 80	1 1:00 0	10:00 20:00 0.1	F2
ST-61&48	*gln*A4 (18, 202)				Same as ST-21 (*gln*A1)						
	*unc*A5 (189) *unc*A17 (336)				95 61 72	95 40 80	20:00 20:00 20:00	95 40 80	1 2:00 0	20:00 20:00 0.1	F2, F3

Table 7.3 Primer and probe sequences for SNP assays based on the Lightcycler

Assay for clonal complex	Alleles and position of SNP(s)	Type of Assay	Primer sequences	Probe sequences	T_m
ST-21	***gln*A1 108,267**	duplex	**ST21*gln*R** GGATCAGGCGTAAAAGG **ST21*gln*F** AACCCTTGAAAAAGTAGGTC	**ST21*gln*108S** LCred-640-GCCACTATTTTTAAGGTGTTCTATTGCT-phos **ST21*gln*108A** TCTGGTCCAAAGTAAGCAGTATCAGCT-fluo **ST21*gln*267S** LCred-705-ATCGGTAAATTCTCTATCATCATTCCACT-phos **ST21*gln*267A** TGTTTCTTGGCCTGTGTCCAGTATTGTAG-fluo	53
	***tkt*_1 330**	uniplex	**ST21*tkt*R** CCATCTCCGCAAAGACA **ST21*tkt*F** AGCACAAGGATTTGAAGT	**ST21*tkt*330S** LCRed-640 ATAGAGATATTGTTGCTATCATAAATAAGTATGAAG TTATCA-Phos **ST21*tkt*330A** CGTTAAAGGCTAAACCTACATCGCCTT-Fluo	51
ST-45	***glt*A10 201, 225**	duplex	**ST45*glt*10F** ACTAAGCGATGAAACAGC **ST45*glt*10R** CCTTACGATCATGTAGAG	**ST45*glt*201A** ATGTAAAAGAAGAATATATGGAAATGGCAGCTAGAATAG-Fluo **ST45*glt*201S** LCred-705 GCTAAAATCCCTACTATAGTCGCCAC-Phos **ST45*glt*225A** TTTCTGTAAAACCACGATCTAAATTTGGATAAGCCA-Fluo **ST45*glt*225S** LCred-640 GGAAAGCCGTGTTTATATCTATAAGCG-Phos	50

ST-257	***gly*A62 483**	uniplex	**ST257*gly*483F** GGGCCACGCAAAGTTTTA **ST257*gly*483R** ATGCAAAACAAGTAAGAACCA	**ST257*gly*483A** ATGTAATTGCTGCAAAAGCAGTAGGATTTAAATT-Fluo **ST257*gly*483S** LCred-640 ATCTTAGCGATGAGTGGAAAGTCT-Phos	50
ST-61	***gln*A4 18, 202**	duplex	***gln*4-17202R** ACTTTGAGCATCAGGTTT ***gln*4-17202F** AATACTGGACACAGGCC	***gln*4-202A** TCATCATTCCACTCTCCTTCTTCGGT-Fluo ***gln*4-202S** LCred201S CAACTTCATACTTAGAACAATGAGCAGT-Phos ***gln*4-18A** CATTTGTCCTTTGTAAATATCATACACATCACAAAATACTATG-Fluo ***gln*4-18S** LCred-705 AGTAGGGTCAGCTGTAAAAGGAT-Phos	53
	***unc*A17 336**	duplex with *unc*A5	**ST61*unc*336F** TTAGGTGCGGGTTCTT **ST61*unc*336R** ACTTGTTTAGTTGCTTTAATTTGAG	**ST61*unc*336A** CGGGCGATGTTTCAGCTTATATTCCAA-Fluo **ST61*unc*336S** LCred 705-ATGTTATTTCGATCACTGATGGACAAA-Phos	40
ST-48	***gln*A4 18, 202**	duplex	Same as above		53
	***unc*A5 186,189**	duplex with *unc*A17	**ST48*unc*189F** TTAGAGACAACGCAAAACAT **ST48*unc*189R** ATGCCGTCAAAGAACC	**ST48*unc*189A** CGTCGTCCTCCAGGTCGTGA-Fluo **ST48*unc*189S** LCred-640 CTTATCCAGGTGATGTTTTTTACCTTCAT-Phos	40

F -forward primer, **R**-reverse primer, **A**-anchor probe, **S**-sensor probe, Lcred-640/705 –sensor dye, Phos- phosphate, Fluo-fluorescein

7.3.2 SNP assay for clonal complex ST-21

(i) Design of primers and probes

The alleles, *glnA1* and *tkt*_1 were identified as suitable representative alleles for the clonal complex ST-21. Within these alleles the associated SNPs (identified in Chapter 6) are summarised below.

Allele	SNP
glnA1	**A→ G (108bp)**
	C→T/A (267bp)
tkt_1	**T→C (330bp)**

Two assays were designed utilising the LC probe design software, (i) a duplex assay for detection of the two SNPs within the *glnA1* allele, and (ii) a separate uniplex assay to detect the one SNP contained within the *tkt*_1 allele. For detection of each SNP two probes were required, the sensor probe which lay over the region of the SNP and was labelled with a fluorescent dye, either LC red 640 or LC red 705 and an anchor probe which lay 1-3 bases away from the sensor probe and was labelled with fluorescein. For the *glnA1* assay one set of forward (4-20bp) and reverse (360-385bp) primers were designed, also one set of internal probes (anchor + sensor probe LCred 640) to detect the SNP at 108bp and a second set of internal probes (anchor + sensor probe LCred 705) to detect the SNP as 267bp (The primer and probe arrangement is shown in Figure 7.1)

For the *tkt*_1 assay forward (217-233bp) and reverse (384-401bp) primers were designed and one set of internal sensor (labelled LC red 640) and anchor probes. The primer and probe arrangement is shown in Figure 7.2

Figure 7.1 Schematic showing positioning of primers and probes for clonal complex ST-21 assay with SNPs at 108bp and 267bp based on the allele *gln*A1

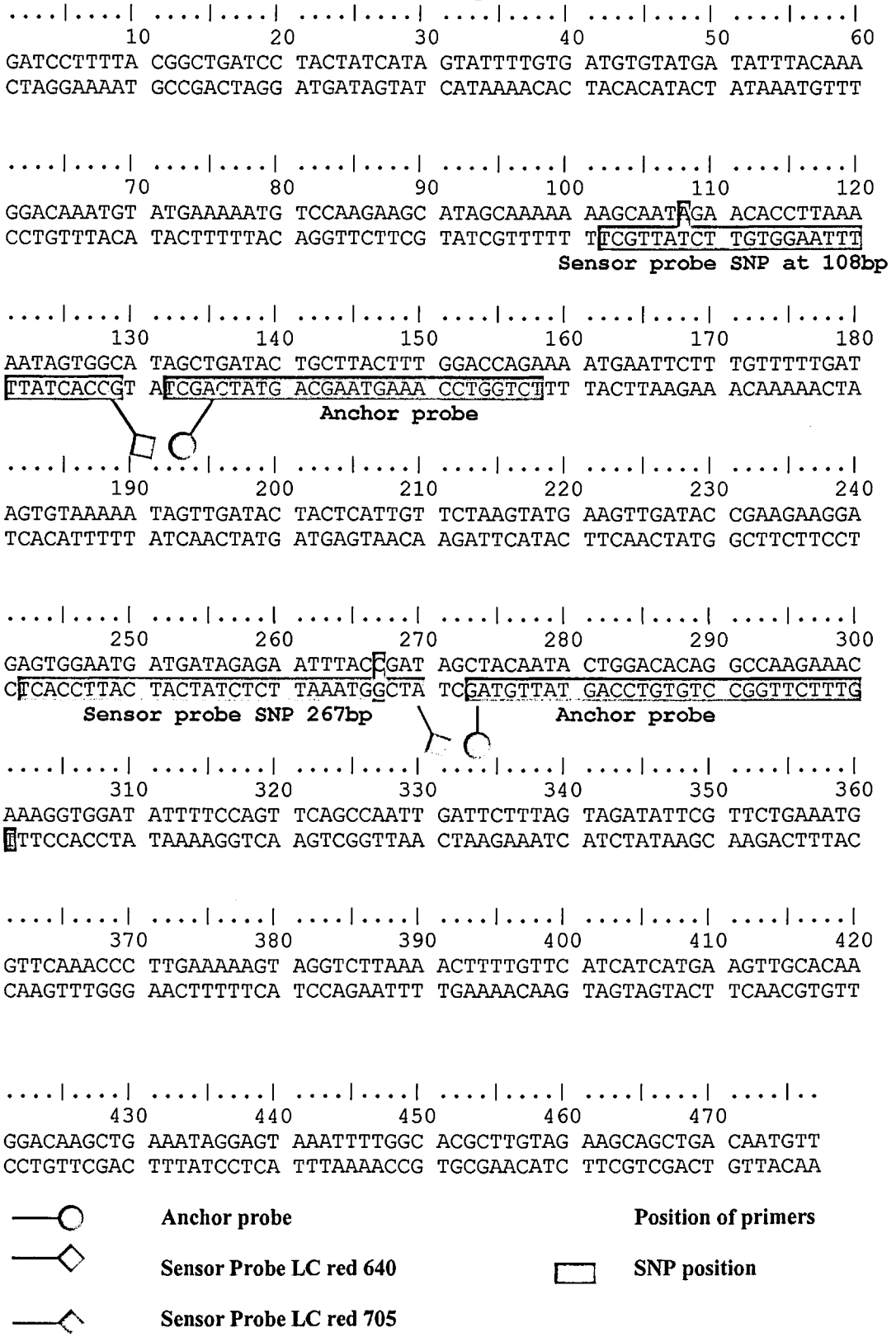
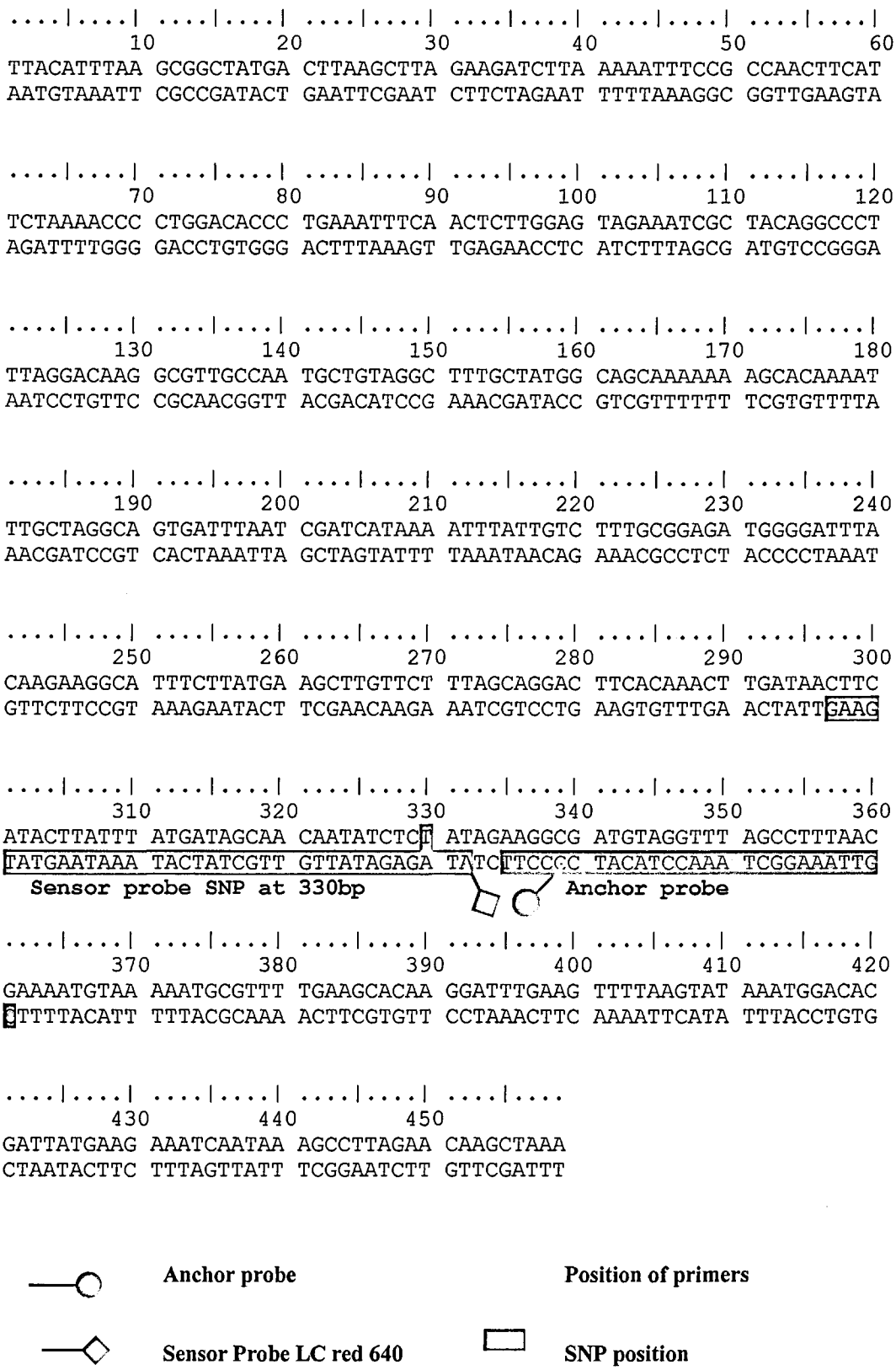


Figure 7.2 Schematic showing positioning of primers and probes for ST-21 assay SNP at 330bp based on the allele *tkt*1



(ii) Melting Curve Analyses for MLST Reference Isolates

For the SNP at 108bp, three main different melting peaks were seen (Figure 7.3a). The highest temperature, indicative of a perfect match between probe and target sequence was indicated by peaks with a melting temperature of 68°C. Mismatched probe/target sequence combinations melted at lower temperatures showing peaks at 66°C or 62°C. The highest melting temperature was only seen in the two isolates of the reference collection, isolates 13254 (ST-21) and 13259 (ST-49) due to the presence of an A at this position (Table 7.4). The other reference isolates showed lower melting peaks, with the temperature of 66°C representing alleles with a G at the same position. In addition a melting temperature of 62°C was seen for some reference isolates, due to the additional SNP at 112bp (T) (Appendix 2.1).

The other *glnA1* SNP at 267bp (Figure 7.3b) had a high melting peak at 68°C. Mismatched probe/target sequences showed peaks at temperatures of 64°C. The highest melting temperature of 68°C was seen in eleven of the fifteen reference isolates indicating the presence of the C at 267bp in the *gln*A allele, which was not just confined to the *gln*A1 allele (Table 7.4) A lower melting temperature of 64°C was seen in isolates 13261 (ST-61), 13258 (ST-48), 13255 (ST-22) and 13268 (ST-403) where the alleles contained a T at this position (Appendix 2.2)

For the assay based on the allele *tkt*_1 (Figure 7.3c) the highest melting temperature was at 66°C, with mismatched probes showing melting peaks at the lower temperatures of 60-64°C. The highest melting temperature was seen in three of the reference isolates 13254 (ST-21), 13263 (ST-206) and 13258 (ST-48), where the T at 330bp on the *tkt*_1 allele was present. The lower temperature of 64°C was due to the

base change to C at this position. The lower temperatures of 62°C and 60°C were seen in isolates 13264 (ST-257) 13262 (ST-177) and 13267 (ST-353) were indicative of further base changes within the 305-330bp region of the probes (Figure 7.3c).

When all three SNPs were detected, indicated by a highest temperature (°C) of the melting peak for each, 68(*glnA*108), 68(*glnA*267), and 66(*tkt*_330) the isolate was deemed to belong to clonal complex ST-21. When applied to the 15 reference isolates, results were 100% concordant with only the 1 isolate, 13254 (ST-21) assigned to the ST-21 clonal complex (Table 7.4).

Throughout the range of assays described here it was impossible to identify SNPs, which were entirely diagnostic for the key allele. For example, it would have required up to five SNPs for the identification of the *gln*A1 allele alone. Although this would have provided exact confirmation of the allele, the time and effort taken to identify all the SNPs would vastly outweigh any benefits of a rapid SNP approach. Specificity was determined to be sufficient for the reference isolates by choosing SNPs, which could be used for distinction of the predictive allele with the least possible number of other alleles. For example within the ST-21 assay the SNPs used (*gln*A1 108 and 267) were not exclusively found within this complex. A highest temperature of 68°C for both these SNPs was seen in the clonal complex ST-49 isolate. This was compensated for by use of the multi SNP approach in adopting more than one allele, *gln*A1 and *tkt*_1 for distinguishing the clonal complex ST-21.

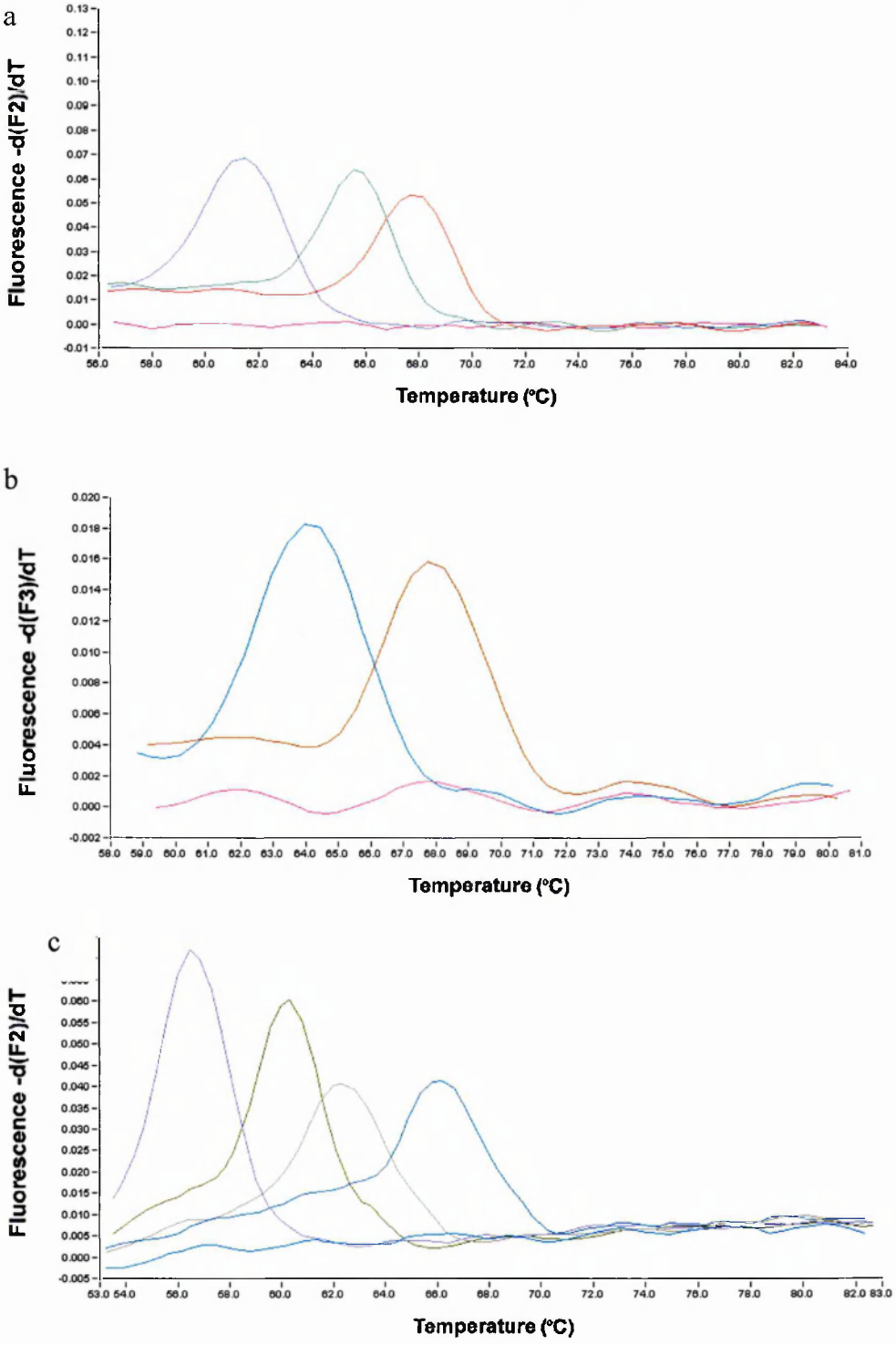


Figure 7.3 a,b & c Example of ST-21 Clonal Complex SNP Melting Curves.

Hybridisation probe fluorescent data showing negative first derivative plots of melting peaks for each SNP for clonal complex ST-21. (c&d duplex reaction, e single reaction). (a) SNP at *gln*A1 108, highest T_m 68. (b) SNP at *gln*A1 267, highest T_m 68. (c) SNP *tkt*_1 330, highest T_m 66.

(iii) Application to larger sample numbers

Following evaluation and optimisation of each assay using the reference isolates, the assays were applied to a panel of 221 diverse samples from retail chicken, human enteritis and animals, which had been characterised by MLST. For this panel of isolates, 51 had been identified by full MLST to belong to ST-21 clonal complex and consistent results were obtained with the SNP strategy (Table 7.5). The same highest melting temperature profiles were seen at each SNP for all of the 51 ST-21 clonal complex assigned isolates, confirming the validity and reproducibility of the assay based upon melting peak data.

The large and diverse nature of this clonal complex, suggested that it would be necessary to use two predictive alleles for accurate identification, which has been confirmed in the results here. Four isolates within other clonal complexes (ST-42, ST-257, ST-49 and ST-22) possessed the *gln*A1 allele but not the *tkt*_1 allele (Table 7.6). It would be highly probable to identify isolates assigned into the ST-49 clonal complex, based on the presence of only the *gln*A1 allele.

Table 7.5 Results for all ST-21 assigned isolates (n=51) by full MLST and consistent results by SNP analyses using T_m results.

Strain	Source	Allelic Profile							ST	SNP T_m result (°C)		
										*gln*A1 108	*gln*A1 267	*tkt*_1 330
47762	chicken offal or meat	4	1	1	3	1	1	5	712	68	68	66
47771	chicken offal or meat	2	1	1	3	2	1	5	21	68	68	66
47785	chicken offal or meat	4	1	1	3	4	1	5	107	68	68	66
47860	human stool	2	1	21	3	2	1	5	53	68	68	66
48237	human stool	2	1	1	3	2	1	5	21	68	68	66
48279	human stool	2	1	21	3	2	1	5	53	68	68	66
48398	chicken offal or meat	2	1	5	3	2	1	5	19	68	68	66
48474	chicken offal or meat	2	1	5	3	2	1	5	19	68	68	66
48479	chicken offal or meat	2	1	6	3	4	1	1	722	68	68	66
48540	human stool	2	1	21	3	2	1	5	53	68	68	66
48710	human stool	2	1	21	3	2	1	5	53	68	68	66
48825	human stool	2	1	1	3	7	1	5	104	68	68	66
48832	human stool	2	1	1	3	7	1	5	104	68	68	66
49006	human stool	2	1	21	3	2	1	5	53	68	68	66
49390	human stool	2	1	1	3	2	1	5	21	68	68	66
49428	human stool	2	1	1	3	2	1	5	21	68	68	66
49438	human stool	2	1	1	3	2	1	5	21	68	68	66
49546	human stool	2	1	21	3	2	1	5	53	68	68	66
49600	human stool	2	1	5	3	2	1	5	19	68	68	66
49688	human stool	2	1	5	3	2	1	5	19	68	68	66
49828	human stool	2	1	1	3	2	1	5	21	68	68	66
49938	human stool	2	1	5	3	2	1	5	19	68	68	66
50258	chicken offal or meat	8	1	21	3	2	1	5	119	68	68	66
50314	human stool	2	1	1	3	2	1	5	21	68	68	66
50707	human stool	2	1	1	3	7	1	5	104	68	68	66
50785	human stool	8	1	6	3	2	1	1	44	68	68	66
50815	chicken offal or meat	8	1	6	3	2	1	1	44	68	68	66
50963	human stool	2	1	1	3	2	1	5	21	68	68	66
51431	chicken offal or meat	9	1	21	3	2	1	5	347	68	68	66
51689	Human stool	2	1	1	3	2	1	5	21	68	68	66
51763	chicken offal or meat	2	1	1	3	7	1	5	104	68	68	66
51820	chicken offal or meat	8	1	6	3	2	1	1	44	68	68	66
52155	human stool	2	1	1	4	2	1	5	761	68	68	66
52376	human stool	2	1	1	3	2	1	5	21	68	68	66
52498	chicken offal or meat	7	1	1	3	2	1	5	615	68	68	66
52530	chicken offal or meat	8	1	6	3	2	1	1	44	68	68	66
52599	chicken offal or meat	7	1	1	9	7	1	5	742	68	68	66
52724	chicken offal or meat	4	1	1	3	7	1	5	744	68	68	66
52833	human stool	2	1	1	3	2	1	5	21	68	68	66
53211	human stool	2	1	12	3	2	1	5	50	68	68	66
53532	human stool	2	1	1	3	2	1	5	21	68	68	66
53793	Chicken offal or meat	4	1	1	3	2	1	5	640	68	68	66
53798	Chicken offal or meat	8	1	1	62	7	1	5	748	68	68	66
54068	chicken offal or meat	8	1	5	4	2	1	5	752	68	68	66
54349	human stool	2	1	1	3	2	1	5	21	68	68	66

54743	human stool	2	1	1	3	2	1	5	21	68	68	66
54984	human stool	2	1	1	3	2	1	5	21	68	68	66
55053	human stool	2	1	21	3	2	1	5	53	68	68	66
55142	human stool	2	1	21	3	2	1	5	53	68	68	66
11168	human	2	1	5	3	4	1	5	43	68	68	66
13254	beef offal	2	1	1	3	2	1	5	21	68	68	66

Table 7.6 Results for all non ST-21 assigned isolates (n=4) with the *gln*A1 allele. The importance of using the second SNP for this clonal complex is demonstrated.

Strain	Source	Allelic Profile							ST	Clonal Complex
48782	chicken offal or meat	9	1	4	62	4	5	6	286	ST-257
49387	Human stool	1	1	3	4	5	9	3	758	ST-42
50727	Human stool	1	1	6	4	3	3	3	759	ST-22
13259	Human stool	3	1	5	17	11	11	6	49	ST-49

7.3.3 SNP assay for clonal complex ST-45

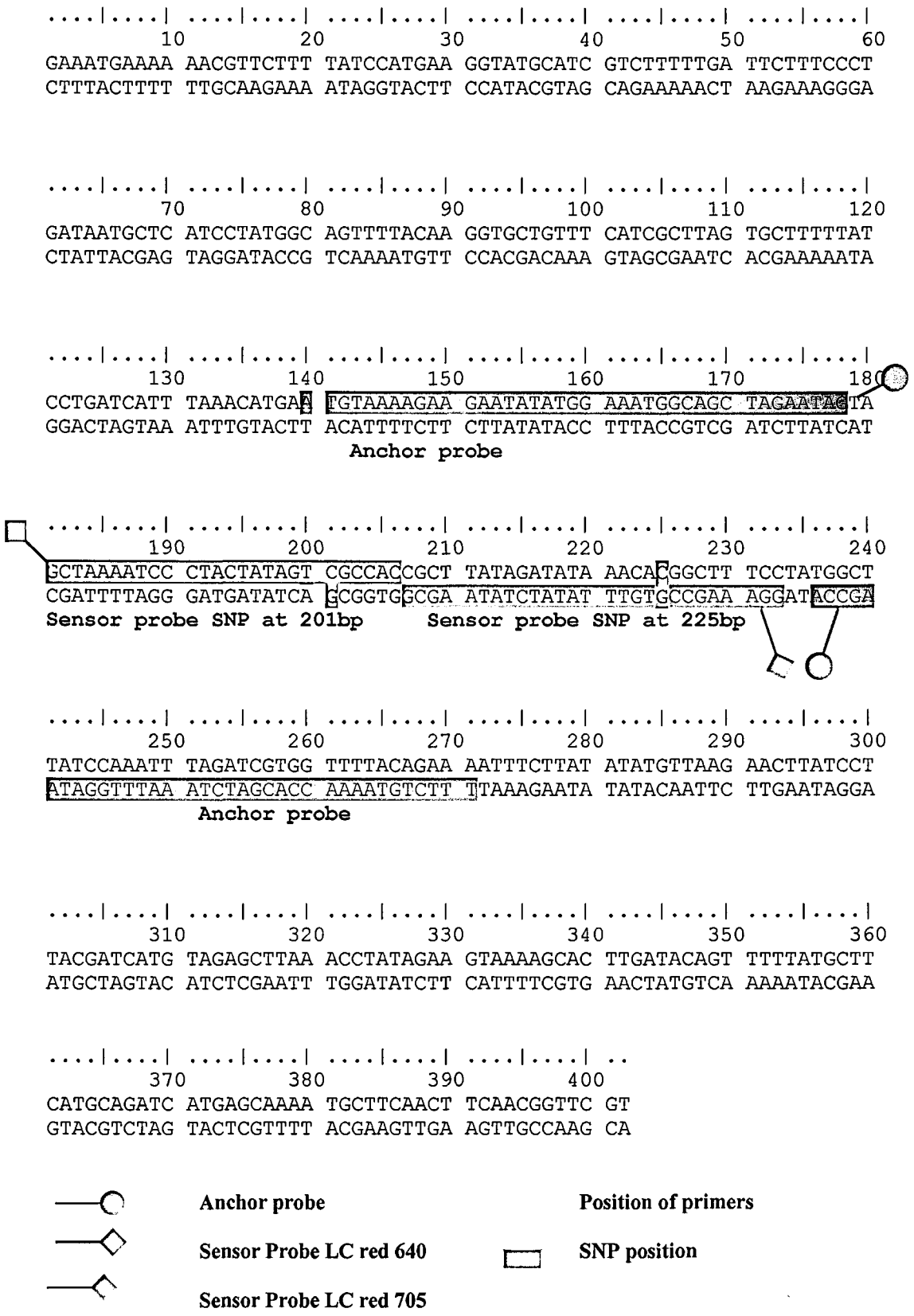
(i) Design of primers and probes

The allele *glt*A10 had been identified in chapter 5 to be a suitable representative allele, and two SNPs were identified at 201bp and 225bp to potentially distinguish the allele *glt*A10 from the remaining *glt*A alleles.

Allele	SNPs
gltA10	C→G/A (201bp)
	C→T (225bp)

One duplex assay was designed utilising the LC probe software to design the primers and one set of probes, the second set were designed manually. The close proximity of the two SNPs required that the probe sets had to be designed back to back, but on different DNA strands. One set of forward (91-112bp) and reverse (298-311bp) primers was designed and two sets of internal probes. For the SNP at 201bp the sensor probe was labelled with LC red 640 and for the SNP at 225bp the sensor probe was labelled with LC red 705 (Figure 7.4).

Figure 7.4 Schematic showing positioning of primers and probes for ST-45 assay, based on allele *gltA*10



(ii) Melting curve analyses for reference isolates

The presence of the SNP at 201bp was confirmed by the highest melting temperature of 66°C. This was seen in isolates 13257 (ST-45) and 13262 (ST-177) due to a C at this position (Table 7.7). Isolate 13262 (ST-177) did not contain the *glt*A10 allele but the *glt*A8 allele, which also contained the C at 201bp. The majority of the other reference isolates showed a melting temperature of 61°C indicative of a G at 201bp. In addition, two reference isolates 13256 (ST-42) and 13264 (ST-257), showed lower temperatures due to the presence of extra SNPs in the area of the probe. These included a further SNP at 200bp on the *glt*A3 allele (13256) and an extra SNP at 204bp on the *glt*A4 allele (13264) (Appendix 2.4)

For the SNP *glt*A10 at 225bp two main melting peaks were obtained. The highest melting peak, indicative of a C at 225bp showed a peak of 65°C, with the lower mismatched probe/target combinations (containing a T) showing peaks at 64°C (Table 7.7). This SNP was more widely distributed among the reference isolates with 9 reference isolates showing the highest melting temperature. The slightly lower melting temperature of 64°C was seen in six of the reference isolates, due to the presence of the T at 225bp or the base change at 207bp.

When both SNPs were detected, indicated by a highest temperature of each melting peak 66°C (*gltA*201) and 65°C (*glt*A225), the clonal complex ST-45 could be identified. This was confirmed with the 15 MLST reference isolates tested, where results were 100% concordant with the one isolate 13257 being identified with the highest melting temperature for both SNPs, therefore assigning to the ST-45 clonal complex (Table 7.7).

(iii) Application to larger sample numbers

For the assay ST-45 only the one allele was used with the detection of two SNPs. Positives (n=30) were indicated as previously described by the profile 66:65°C, which was indicative of the ST-45 isolates. The 225bp SNP was widely distributed across the alleles as a temperature of 65°C was seen for isolates in more than half of the clonal complexes. However, the SNP at 201bp was less widely distributed (Table 7.16).

Within the data set used 30 isolates were identified as ST-45 clonal complex by SNP analysis based on the presence of the SNPs on the *glt*A10 allele, however this result was not consistent with the result obtained by full MLST, where only 28 isolates were identified as ST-45 (Table 7.8). On closer inspection two isolates had been falsely identified as ST-45 due to the presence of the *glt*A10 allele when they were actually assigned to another clonal complex or unassigned (Table 7.9).

This highlighted the importance of using two alleles for correct determination of the clonal complex. If the additional allele *tkt*_7 had been incorporated then increased specificity would be obtained for ST-45 clonal complex assigned isolates. The absence of the allele *tkt*_7 in both of these incorrectly identified isolates would have prevented their identification as ST-45. Observations concerning the implementation of the SNP assays on the lightcycler, throughout this chapter (as discussed later) meant that a further platform was considered to use for SNP detection. The incorporation of an additional allele for this clonal complex to improve the specificity is described in the next chapter.

Table 7.8 Results for all ST-45 assigned isolates (n=28) by full MLST and consistent identification by SNP analyses using T_m results

Strain	Source	Allelic Profile							ST	SNP T_m result (°C)	
										*glt*A10 201	*glt*A10 225
47720	chicken offal or meat	24	7	10	62	42	7	1	756	66	65
47728	chicken offal or meat	2	7	10	4	1	7	1	233	66	65
47732	chicken offal or meat	2	7	10	4	1	7	1	233	66	65
47733	chicken offal or meat	9	7	10	4	1	7	1	320	66	65
48300	chicken offal or meat	33	7	10	4	1	7	1	714	66	65
48311	chicken offal or meat	4	7	10	4	1	7	1	45	66	65
48435	chicken offal or meat	2	7	10	4	1	7	1	233	66	65
48471	chicken offal or meat	2	2	10	4	2	7	1	721	66	65
48885	human stool	4	7	10	4	1	7	1	45	66	65
49294	human stool	4	60	10	4	1	7	1	765	66	65
49594	human stool	4	7	10	4	1	7	1	45	66	65
49979	human stool	4	17	10	4	1	7	1	766	66	65
50254	chicken offal or meat	4	4	10	4	2	7	1	734	66	65
50313	human stool	4	7	10	4	1	7	1	45	66	65
50725	human stool	4	7	10	4	1	7	1	45	66	65
51009	human stool	4	7	10	4	1	7	1	45	66	65
51405	chicken offal or meat	2	7	10	4	1	7	1	233	66	65
51797*	chicken offal or meat	9	1	10	3	2	1	5	368	66	65
51953	human stool	4	7	10	4	1	7	5	241	66	65
53892	human stool	4	2	10	4	1	7	1	782	66	65
53967	chicken offal or meat	9	7	10	2	1	7	1	750	66	65
54401	human stool	4	7	10	1	1	7	1	25	66	65
54879	human stool	4	7	10	4	1	7	1	45	66	65
56050	chicken offal or meat	2	7	10	1	1	7	1	755	66	65
56052	chicken offal or meat	8	7	10	4	1	7	1	754	66	65
29615	wild bird	4	7	10	4	1	7	1	45	66	65
29645	Cattle	4	7	10	4	1	7	1	45	66	65
29693*	ovine abortion	2	2	10	4	7	71	1	726	66	65
29703	canine	4	7	10	4	1	7	5	241	66	65
13257	human stool	4	7	10	4	1	7	1	45	66	65

*** Highlighted samples identified as ST-45, due to presence of *glt*A10 allele. Both samples should be assigned to different clonal complexes. (51797*- ST-21 and 29693*-unassigned)**

Table 7.9 Results for all non ST-45 assigned isolates (n=2)

Clonal Complex by SNP strategy	Allelic Profile	ST	Clonal Complex by full MLST
ST-45	2, 2, 10, 4, 7, 71, 1	726	Unassigned
ST-45	9, 1, 10, 3, 2, 1, 5	368	ST-21

7.3.4 SNP assay for clonal complexes ST-48 and ST-61

(i) Design of primers and probes

The relatedness of these two complexes, shown by their sharing of common representative alleles, allowed a combined strategy to be designed. The common allele used was the *gln*A4 allele for the identification of both complexes, then subsequent further alleles, (*unc*A17 for ST-61 and *unc*A5 for ST-48) for identification of each individual clonal complex. The SNPs used for detection of these alleles are shown below.

Allele	SNPs
***gln*A4**	**C→T** (18bp)
	G→A (202bp)
***unc*A17**	**C→T** (336bp)
***unc*A5**	**A→T** (186bp) **T→C/A** (189bp)

Two duplex assays were designed utilising the LC probe design software (i) a duplex assay for the detection of the two SNPs within the *gln*A4 allele, and (ii) a further duplex assay for the detection of the SNPs within the *unc*A5 and *unc*A17 alleles.

The most specific SNPs to distinguish the *gln*A4 allele were located at 18bp and 202bp. The SNP at 202bp was easily identifiable and probes were successfully designed to detect the SNP. However, for the SNP at 18bp, the process was more difficult due to the close proximity of the SNP to the beginning of the allele sequence, meaning there was insufficient room to fit in a primer and probe sequence. To facilitate the use of this SNP, the allele sequence was extended by 50 bases at the 5' end, so that a primer could be designed on the allele before the location of the SNP. The extra sequence for *gln*A4 in this region was obtained by the alignment of

*gln*A4 alleles before trimming to size to determine any variability within the area (Figure 7.5).

For the allele *unc*A5 (ST-48), forward (71-90bp) and reverse (251-271bp) primers were designed and a set of internal hybridisation probes using the fluorescent extension LC red 640 (Figure 7.6). As the ST-61 reaction was also based on the *unc*A locus, it was possible to multiplex these two assays together to create a more convenient method for distinguishing the two clonal complexes. Forward (247-262bp) and reverse (434-458bp) primers were designed as well as internal probes; the sensor probe was labelled with LCred 705 to allow for successful multiplexing (Figure 7.7).

Figure 7.5 Schematic showing positions of primers and probes for *gln*A4 Assay.

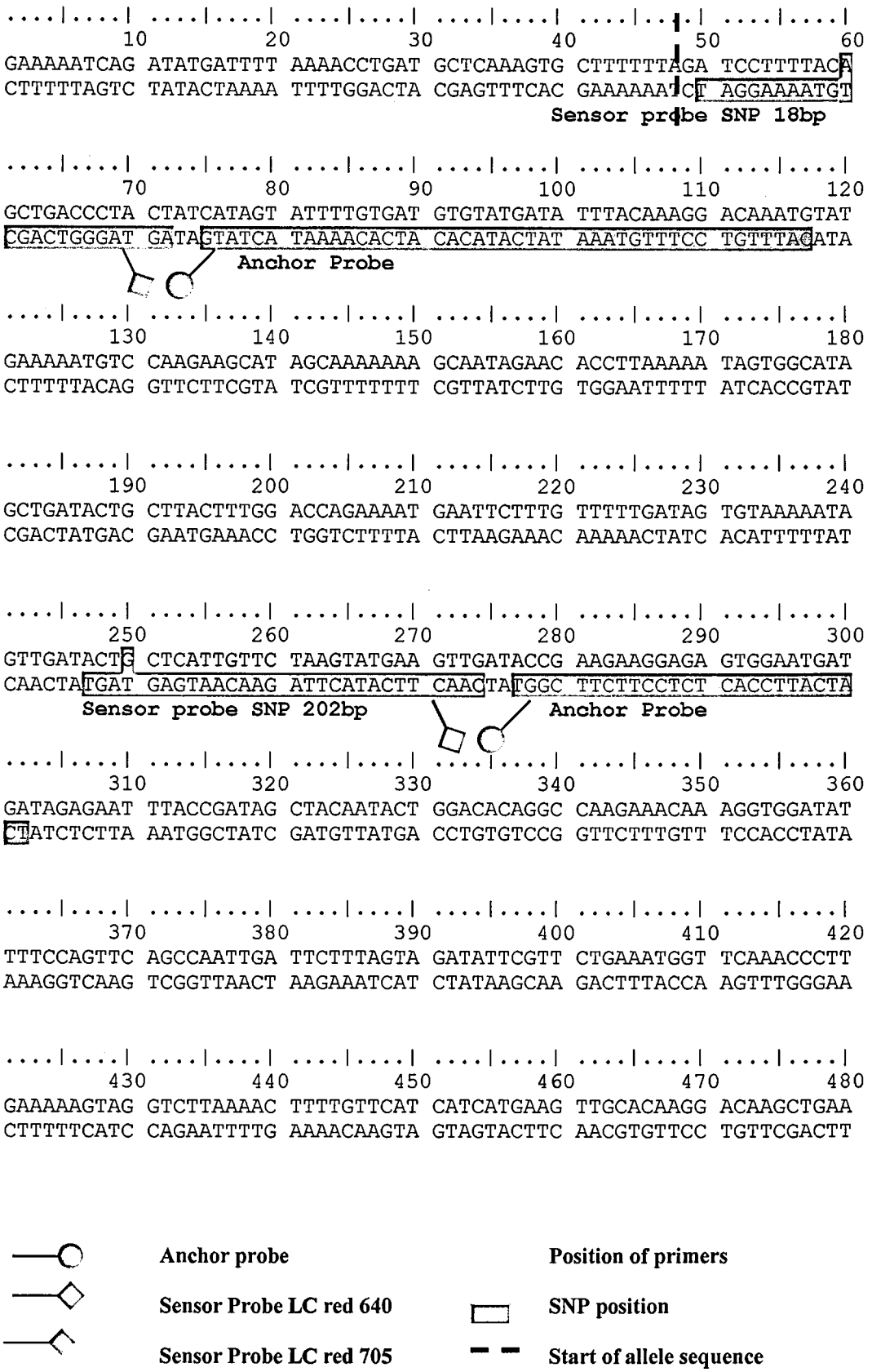


Figure 7.6 Diagram showing positions of primers and probes for ST-48 assay based on the allele *unc*A5.

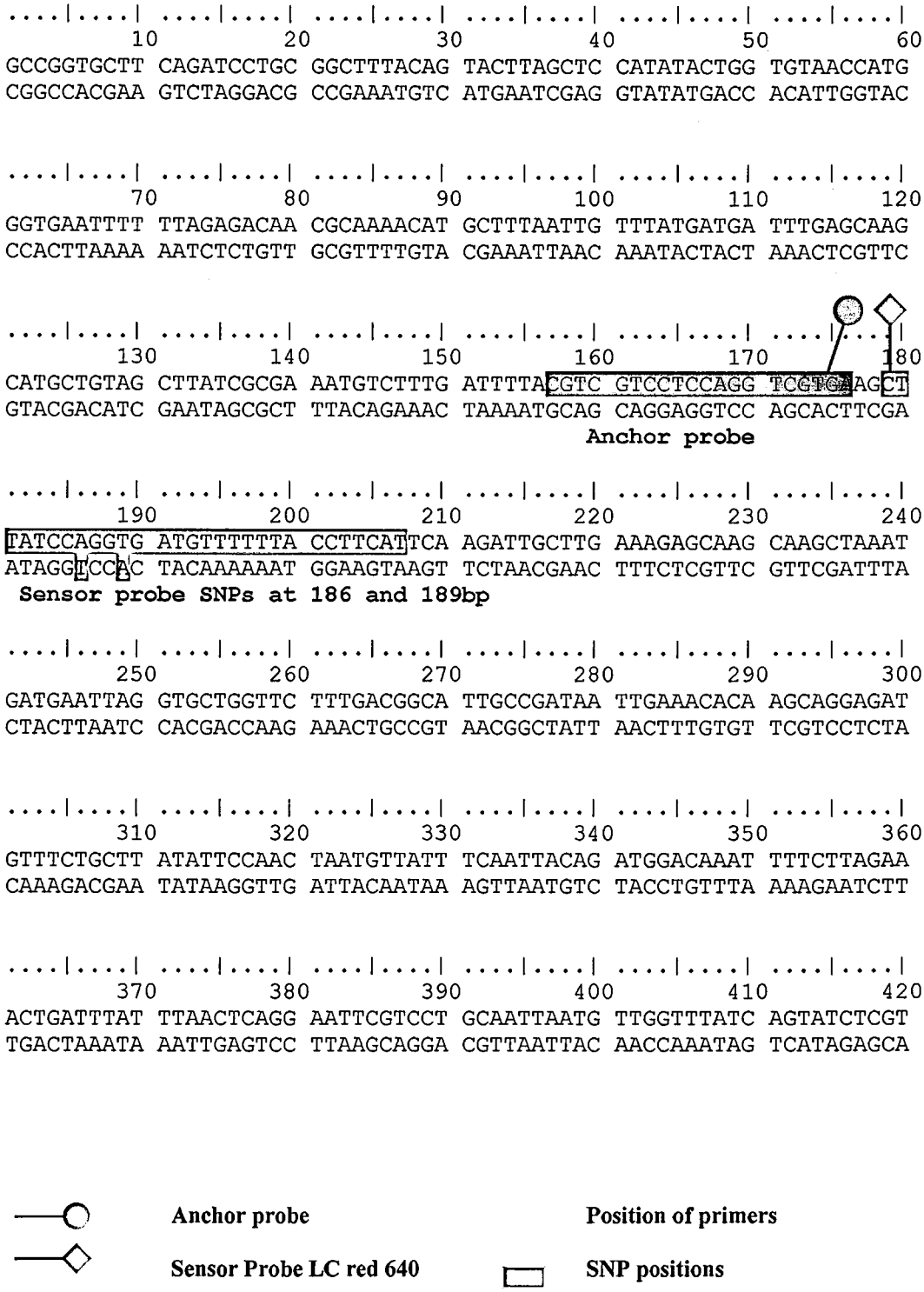
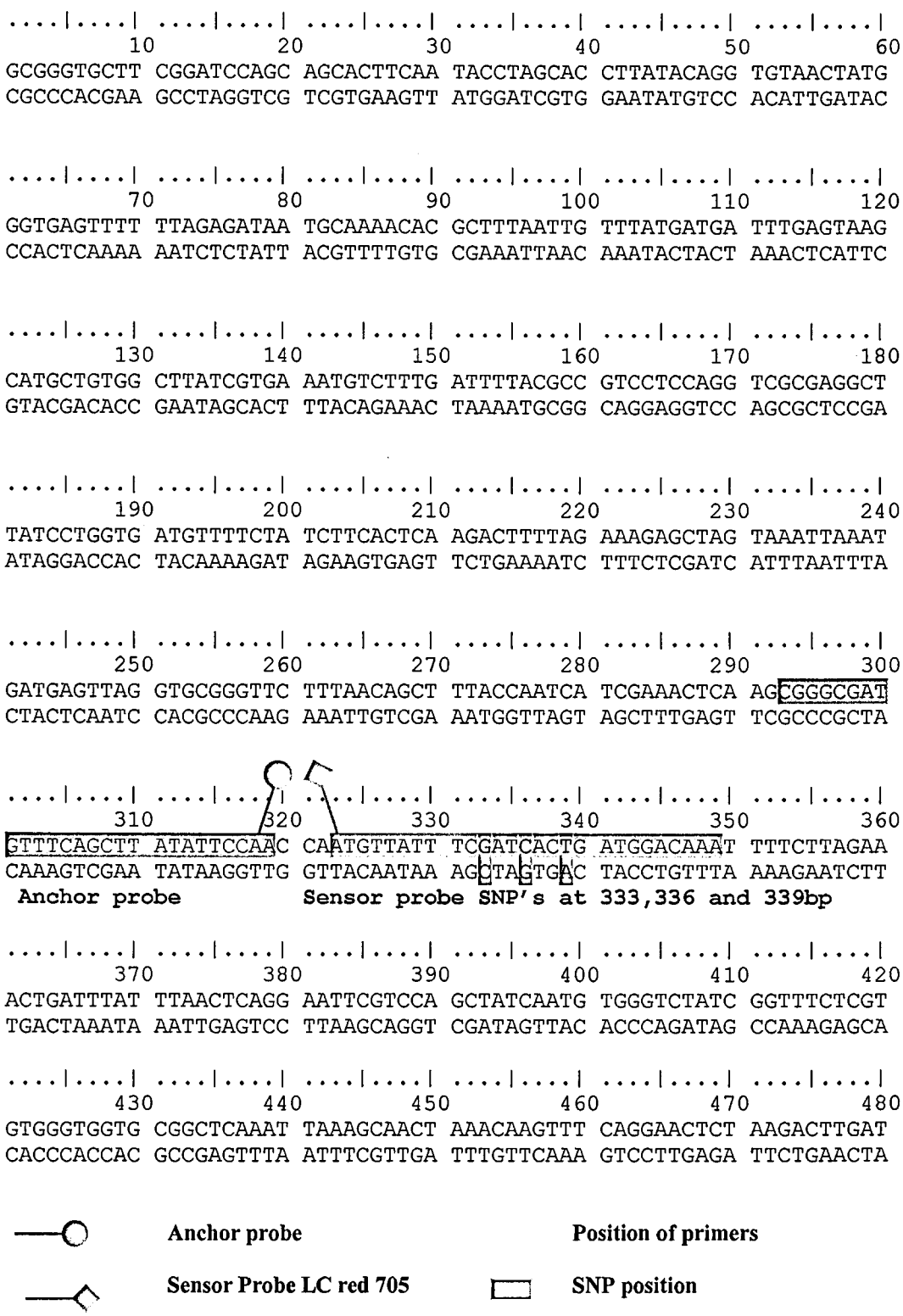


Figure 7.7 Diagram showing positions of primers and probes for ST-61 assay, based on allele *unc*17



(ii) Melting curve analyses for reference isolates

For the SNP *gln*A4 at 18bp a range of different melting peaks were seen. The highest temperature peak of 65°C, indicative of the perfectly matching probe containing the SNP (C) was present in isolates 13261 (ST-61), 13258 (ST-48), 13268 (ST-403) and 13263 (ST-206) (Table 7.10) A lower temperature peak of 61°C in 13255 (ST-22) and 13257 (ST-45) corresponded to the presence of a T at this position. A further temperature peak of 58°C was seen in the remainder of the reference isolates. This corresponded to the presence of the T at 18bp as well as a G at position 12bp (Appendix 2.6).

For the SNP at 202bp the presence of a G was confirmed by the highest temperature of 65°C which was present in isolates 13261 (ST-61), 13258 (ST-48) 13265 (ST-354) and 13255 (ST-22) (Table 7.10). A lower temperature peak was seen at 60°C corresponding to an A in this position in the remainder of the reference isolates. Only with the highest profile for both these SNPs could the presence of the allele *gln*A4 be confirmed. This was only present in isolates 13261 (ST-61) and 13258 (ST-48) (Appendix 2.7).

To further identify the clonal complex ST-61 and ST-48 assigned reference isolates, the second duplex assay based on the *unc*A allele was required. With the *unc*A5 SNPs (for ST-48) at 186 and 189bp, two different sized peaks were observed. The highest temperature indicative of the perfect match, and the presence of an A and T, showed a melting peak at 65°C which were present in four reference isolates, all with

the *unc*A5 allele, isolates 13254 (ST-21), 13263 (ST-206), 13258 (ST-48) and 13268 (ST-403) (Table 7.10).

For the other component of the assay, based on the allele *unc*A17, the highest peak, indicative of this allele was seen at 65°C. Due to this allele having a number of SNPs located together (333bp, 336bp and 339bp) in the area used for the probe. This explained the large shift in melting temperature seen between the presence of the SNPs within *unc*A17 (65°C) only present in isolate 13261 (ST-61) (Table 7.10) and the base changes present in the remaining isolates (Appendix 2.9), where the shift in melting temperature was 15°C, with negative samples showing melting peaks of around 50°C.

For the correct identification of these complexes, two different highest melting temperature profiles were required, as shown:

Clonal complex	**Required melting temperature (°C)**			
	***gln*A4**		***unc*A5**	***unc*A17**
	18bp	**201bp**	**190bp**	**336bp**
For ST-61	65	65	Less than 65	65
For ST-48	65	65	65	Less than 65

When tested with the 15 reference isolates results were 100% concordant with only the two isolates 13261 (ST-61) and 13258 (ST-48) showing the correct highest temperature.

(iii) Application to larger sample numbers

Using the combined approach in identifying the clonal complexes ST-61 and ST-48 the complete set of ST-48 isolates (n=28) and the ST-61 isolates (n=6) were identified in the dataset (Tables 7.11 and 7.12). The *gln*A4 assay was successful in identifying the ST-61 and ST-48 isolates although it did also identify some of the isolates within the ST-206 clonal complex. On investigation 20% of isolates assigned to the ST-206 clonal complex actually possessed the allele *gln*A4, although in conjunction with the extra specific SNPs used for the identification of ST-61 and ST-48, the specificity for this data set was not compromised.

For *unc*A5 the highest temperature of 65°C was also seen within other clonal complexes ST-21 and ST-206. However in conjunction with the allele *gln*A4 only the ST-48 isolates were identified (Tables 7.11 and 7.16). For ST-61 clonal complex the SNPs for the *unc*A17 allele were highly discriminatory and the greatest shift in the melting temperature from 65°C (positive for *unc*A17) to 50°C (negative for *unc*A17) occurred as seen for the reference isolates. The combined approach for the identification of these two clonal complexes offered a more rapid approach, where only two lightcycler runs were performed, taking two hours (after time for 1st round PCR step) rather than the three hours of Lightcyler cycling which would be required if the two *unc*A alleles were in separate reactions.

Table 7.11 Results for all ST-48 assigned isolates (n=21) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP T_m result (°C)			
										*gln*A4 18	*gln*A4 202	*unc*A5 189	*unc*A17 336
48538	human stool	2	4	1	2	7	51	5	205	65	65	65	50
49009	human stool	2	4	1	2	7	1	5	48	65	65	65	50
49474	human stool	2	4	1	2	7	1	5	48	65	65	65	50
50944	human stool	2	4	1	2	7	51	5	205	65	65	65	50
51531	human stool	2	4	1	2	7	51	5	205	65	65	65	50
51926	human stool	2	4	2	2	6	1	5	38	65	65	65	50
52142	human stool	2	4	1	2	7	1	5	48	65	65	65	50
52477	chicken offal or meat	9	4	1	2	7	1	5	738	65	65	65	50
52484	chicken offal or meat	9	4	1	2	7	1	5	738	65	65	65	50
52834	human stool	2	4	1	2	7	1	5	48	65	65	65	50
53086	chicken offal or meat	8	4	1	2	7	1	5	414	65	65	65	50
53703	chicken offal or meat	4	4	1	2	7	1	5	739	65	65	65	50
53706	chicken offal or meat	4	4	1	2	7	1	5	739	65	65	65	50
53785	chicken offal or meat	4	4	1	3	7	1	5	749	65	65	65	50
53810	chicken offal or meat	4	4	1	3	7	1	5	749	65	65	65	50
53854	human stool	2	4	1	2	7	51	5	205	65	65	65	50
54720	human stool	2	4	1	2	7	1	5	48	65	65	65	50
29697	ovine abortion	2	4	2	2	6	1	5	38	65	65	65	50
76792	human OB	2	4	5	2	7	1	5	66	65	65	65	50
76879	human OB	2	4	1	2	7	1	5	48	65	65	65	50
13258	lamb offal	2	4	1	2	7	1	5	48	65	65	65	50

Table 7.12 Results for all ST-61 assigned isolates (n=6) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP T_m result (°C)			
										*gln*A4 18	*gln*A4 202	*unc*A5 189	*unc*A17 336
29644	cattle	1	4	2	2	22	9	17	620	65	65	63	65
29662	cattle	1	4	2	2	6	9	17	620	65	65	63	65
13261	beef offal	1	4	2	2	6	3	17	61	65	65	63	65
55706	cattle	1	4	2	16	6	3	17	60	65	65	63	65
55729	cattle	1	4	19	2	6	3	17	36	65	65	63	65
56054	cattle	2	4	2	2	6	4	17	81	65	65	63	65

7.3.5 SNP assay for clonal complex ST-257

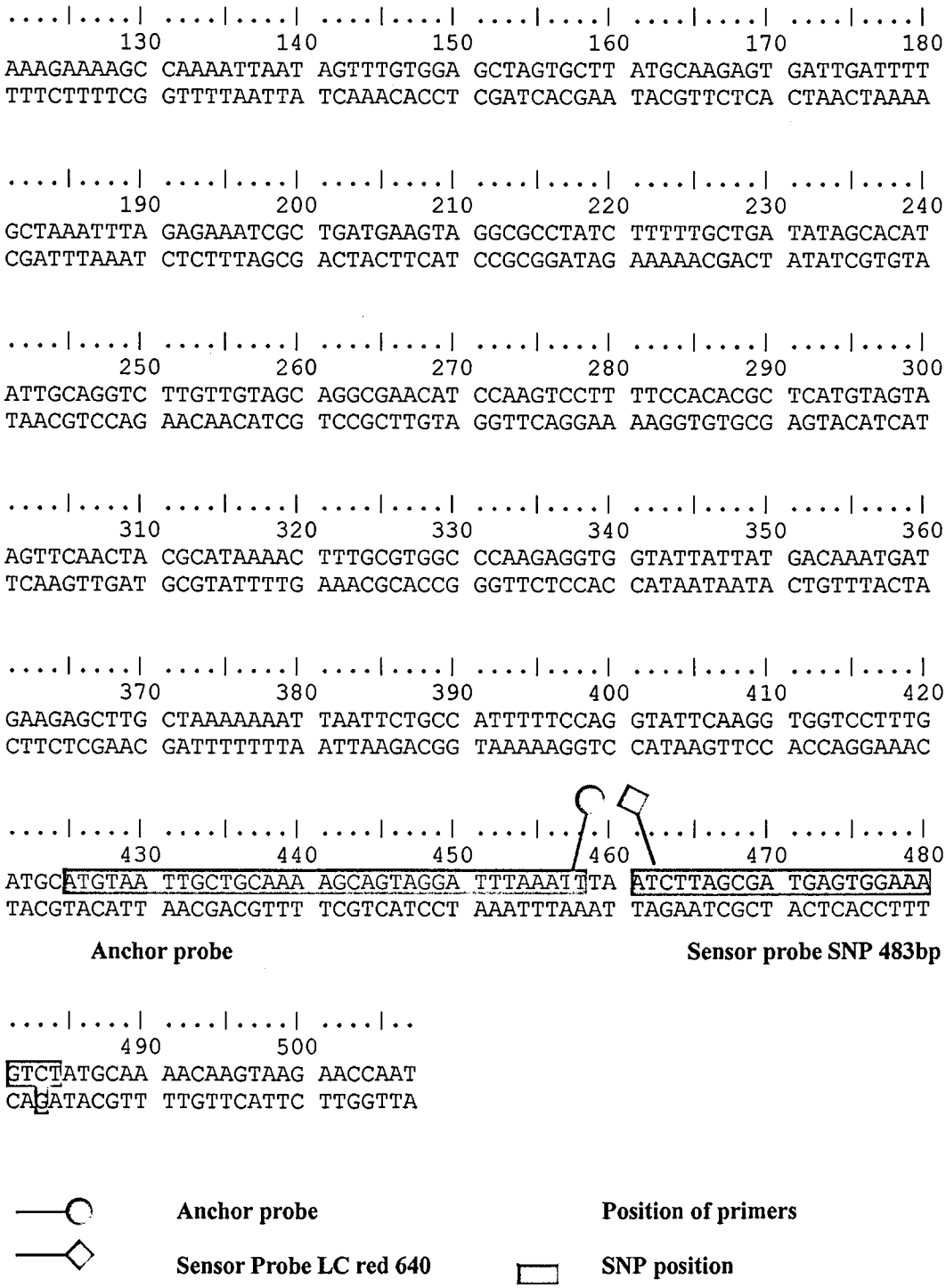
(i) Design of primers and probes

The allele *glyA*62 was identified as a suitable representative allele for the clonal complex ST-257. The SNP identified was located at 483bp (Chapter 6).

Allele	SNPs
glyA62	C→T (483bp)

A uniplex assay was designed to detect the SNP. As in the case of ST-45, the location of the SNP towards the end of the allele sequence made it difficult to design primers and probes which successfully met the parameters of the software. This could only be achieved by the placement of the sensor probe upstream from the SNP site and placing the SNP in the penultimate base of the sensor probe. The forward (315-332bp) and reverse (485-507bp) primers were designed with the probes internal to these. The sensor probe was labelled with LC red 640 (Figure 7.8).

Figure 7.8 Schematic showing positions of primers and probes for ST-257 assay, based on the allele *glyA* 62



(ii) Melting curve analyses for reference isolates

With the SNP *glyA*62 at 483bp two main melting peaks were seen. The highest temperature, indicative of a perfect match between the probe and the target SNP (C) was indicated by a peak at 66°C, the other alleles containing a T showed a peak with melting temperature of 61°C.

Application of the assay to the 15 reference isolates successfully identified the clonal complex ST-257 isolate 13264 by the highest melting temperature of 66°C (Table 7.13). All of the remaining reference isolates showed the lower melting temperature of 61°C due the presence of a T at this position. No variation in this melting temperature was seen due to the location of this SNP in a highly conserved portion of the gene and due to no other sequence variation in the region of the probes (Appendix 2.11).

(iii) Application to larger sample numbers

The same distinctive melting peak was seen for the ST-257 isolates throughout the larger sample set. For clonal complex ST-257 a highly distinctive melting temperature was seen for the positive ST-257 isolates (n=29) of 66°C compared to the temperature of 61°C for non ST-257 samples (Table 7.14). It had been determined in Chapter 5 that the allele *gly*A62 was 85% specific for the clonal complex ST-257, therefore it was possible that this allele could occur within other clonal complexes, which has been demonstrated here.

Within the data set used 33 isolates were identified as ST-257 clonal complex by SNP analysis based on the presence of the *gly*A62 allele, however this result was not

consistent with the result obtained by full MLST, where only 29 isolates were identified as ST-257 (Table 7.14). Four isolates had been incorrectly assigned to the ST-257 clonal complex by the presence of the *glyA*62 (Table 7.15). One isolate from clonal complex ST-21 was incorrectly identified, in addition to one isolate from ST-45 and two unassigned isolates.

As in the case for clonal complex ST-45, the importance of using two alleles for correct determination of the clonal complex was demonstrated. If the additional allele *pgm*_4 (from Chapter 5) had been incorporated into the strategy for ST-257 then increased specificity would be obtained and it would be more likely to only identify ST-257 clonal complex assigned isolates. The absence of the other predictive allele *pgm*_4 in these four incorrectly identified isolates would have avoided the detection of these isolates as ST-257. Several observations concerning the implementation of the SNP assays on the Lightcycler, throughout this chapter (as discussed later) meant that a further platform was considered for SNP detection. The incorporation of a further allele for this clonal complex to improve the specificity is described in the next chapter.

Table 7.14 Results for all ST-257 assigned isolates (n=29) by full MLST and consistent identification by SNP analyses using T_m results

strain	source	Allelic Profile							ST	SNP T_m result (°C)
										*gly*A62 483
47720*	chicken offal or meat	24	7	10	62	42	7	1	756	66
47781	chicken offal or meat	2	2	4	62	4	5	6	367	66
47782	chicken offal or meat	2	2	4	62	4	5	6	367	66
48298	chicken offal or meat	9	2	4	62	4	5	6	257	66
48323	chicken offal or meat	9	2	5	62	4	5	1	717	66
48352	chicken offal or meat	9	2	4	62	4	5	6	257	66
48384	chicken offal or meat	2	2	4	62	4	5	6	367	66
48393*	chicken offal or meat	2	1	5	62	2	5	6	718	66
48417	chicken offal or meat	2	2	4	62	4	5	6	367	66
48496	chicken offal or meat	4	2	4	62	4	5	6	366	66
48658*	human stool	7	1	6	62	11	67	6	770	66
48675	human stool	9	2	4	62	4	5	6	257	66
48742	chicken offal or meat	2	2	4	62	4	5	6	367	66
48782	chicken offal or meat	9	1	4	62	4	5	6	286	66
49407	human stool	9	2	4	62	4	5	6	257	66
49898	human stool	9	2	4	62	4	5	6	257	66
50820	chicken offal or meat	2	2	4	62	4	5	6	367	66
50830	chicken offal or meat	4	2	4	62	4	5	6	366	66
50995	human stool	9	2	4	62	4	5	6	257	66
50996	human stool	9	2	4	62	4	5	6	257	66
51538	human stool	9	2	4	62	4	5	6	257	66
51790	chicken offal or meat	4	2	4	62	4	5	6	366	66
51817	chicken offal or meat	2	2	4	62	4	5	6	367	66
52514	chicken offal or meat	1	2	4	62	4	5	6	737	66
52788	human stool	9	4	4	62	4	5	6	776	66
53798*	chicken offal or meat	8	1	1	62	7	1	5	748	66
53896	human stool	8	2	5	62	4	5	1	584	66
54309	human stool	9	2	4	62	4	5	6	257	66
54471	human stool	9	10	4	62	4	5	6	777	66
54626	human stool	9	2	4	62	4	5	6	257	66
76665	human OB	9	2	4	62	4	5	6	257	66
12013	Poultry	9	2	4	62	4	1	6	316	66
13264	human stool	9	2	4	62	4	5	6	257	66

Highlighted samples identified as ST-257, due to presence of *gly*A62 allele. Samples* should be assigned to different clonal complexes or unassigned samples

Table 7.15 Results for non ST-257 assigned isolates (n=4) (red- presence of ST-257 predictive allele)

Clonal Complex by SNP strategy	Isolate	Allelic Profile	ST	Clonal Complex by full MLST
ST-257	47720	24, 7, 10, 62, 42, 7, 1	756	ST-45
ST-257	48393	2, 1, 5, 62, 2, 5, 6	718	Unassigned
ST-257	48658	7, 1, 6, 62, 11, 67, 6	770	Unasigned
ST-257	53798	8, 1, 1, 62, 7, 1, 5	748	ST-21

7.3.6 Summary of results for all five clonal complexes

Table 7.16 shows a summary of results for all isolates and the comparison between results obtained by full MLST and with the SNP strategy. Data were consistent for clonal complexes ST-21, ST-61 and ST-48, where equal numbers of isolates were identified by both techniques. For clonal complexes ST-45 and ST-257, the data have shown that greater specificity would be achieved for the accurate identification of these clonal complexes by the addition of a further predictive allele.

Specific Lightcycler melting temperature profiles were obtained, for example, in clonal complex ST-21 the SNP temperature profile of 68:68:66°C was only seen within the ST-21 assigned isolates (Table 7.16). Additionally a different profile was seen for clonal complexes ST-61 and ST-48, where each highest temperature profile of 65:65:65°C was only seen for the specific SNPs. For the other two clonal complexes, the highest temperature profile was not entirely distinct for each clonal complex. For example, the highest temperature for ST-257 of 65°C was seen within three clonal complexes, emphasising the value of increasing the number of loci to improve the resolution.

It was possible to determine the distribution of highest melting temperatures across many clonal complexes (Table 7.16). Only the clonal complex ST-61 had a distinct highest temperature, for the *unc*A17 SNP that was not seen elsewhere in the table. Conversely the highest melting temperature for the ST-45 SNP (*glt*A10 225) of 65°C was widely distributed throughout the Table. It was possible to envisage that based upon the distribution of highest melting temperatures throughout all the clonal complexes, a strategy could be developed for identification of more than the clonal

complexes described here, based upon these existing SNPs. However a great many more isolates would need to be analysed to ensure the data was phylogenetically valid.

7.4 General Discussion

Rapid strain characterisation is critical to the delivery of real time epidemiology to inform the timely recognition of case clusters. Analysis of the allelic variation within the MLST database has allowed the development of a rapid sequence based approach to identify clonal complexes based on detection of SNPs. The characterisation strategy described is based upon the genetic groupings delineated in the internationally accepted web-enabled MLST database and uses standard nomenclature already in place for the characterisation of *C. jejuni* isolates. This means that data obtained can be easily compared with and added to existing databases or epidemiological studies. Furthermore, this approach possesses some of the main advantages of MLST as a typing scheme in that data can be easily compared and transported between laboratories and uses increasingly widely available laboratory equipment.

Using melt curve analyses on the Roche Lightcycler it has been possible to identify SNPs that can identify specific alleles and thus indicate the clonal complex of an unknown isolate. The melting peak analyses are robust and each assay shows an obvious shift in melting temperature between different alleles based on their sequence. Real time allelic discrimination assays have been developed for the clonal complexes ST-21, ST-45, ST-48, ST-61 and ST-257 and have been validated by the MLST reference collection. These assays on the Lightcyler have demonstrated the feasibility of assigning MLST complexes by SNP analysis and suggest that the further development and extension of the scheme would be constructive.

The majority of the isolates from a panel of diverse clinical, environmental and reference strains were correctly assigned to clonal complexes ST-21, ST-45, ST-48, ST-61 and ST-257 as appropriate. The pattern of melting peak profiles was consistent through all the assays. The melting temperature profiles, when compared to full MLST data confirmed that the assays were specific for clonal complexes ST-21, ST-61 and ST-48 when applied to larger sample numbers. The inclusion of additional predictive alleles into the strategy for ST-45 and ST-257 would enable adequate specificity to be obtained for all these clonal complexes.

These assays could permit the rapid identification of *C. jejuni* isolates and preliminary strain identification, which could inform epidemiological investigations. In addition these assays would permit the rapid screening of isolates for specific types of *C. jejuni*. Currently there are no typing methods suitable for rapid detection of specific types, which can be applied for confirmatory sample testing. A strategy whereby specific strains based on their clonal complex could be identified from a patient sample and this subsequently matched with a food sample would be highly desirable. Currently *C. jejuni* typing methods rely upon cultured samples, sometimes difficult to obtain, therefore a DNA based method such as this type of SNP strategy would be advantageous.

The utility and reproducibility of the assays has been confirmed, and the next step, following incorporation of additional predictive alleles, would be to apply them prospectively in larger scale screening studies and for rapid epidemiological investigations. The successful application to isolates from different sources suggests that this approach could be used for characterisation of strains from diverse

backgrounds and the data then related unambiguously to existing MLST data. From this data set 60% of isolates, have been identified by clonal complex, which is consistent with other studies (Dingle *et al.*, 2001). However 40% of isolates were unable to be characterised, due to either being in other clonal complexes, or they were incorrectly assigned. The resolution obtained by full MLST would never be achievable with a SNP approach; therefore this strategy has been designed as a complementary technique to MLST not as a replacement. It could provide a method for elimination of the most common clonal complex assigned isolates, leaving a smaller subset to be analysed by full MLST, therefore reducing costs and analysis time.

A limiting factor of SNP assays is that they are entirely based upon the existing data within the MLST database. The data currently in the PubMLST database are only representative of the isolates that have been both typed by MLST and submitted. It may be somewhat biased in that not all isolates, particularly those of the most common types, will have been submitted. Thus the alleles on which the SNP assays are based may be more specific than actually calculated, due to the circulation of more of the common types within the environment. This success of the strategy so far, and the presence of further predictive alleles and informative SNPs identified in Chapters 5 and 6 would enable the further development of the strategy to incorporate additional clonal complexes.

It was difficult to achieve good quality melting peaks in some assays (e.g. clonal complex ST-21, *tkt*) despite good amplification reactions. In these cases repeat melt curves were carried out on the same samples (up to four times) each time changing

one parameter of the melting step. The main improvement was seen after the extension of the time at the denaturation step in the melt curve process. By extending this from 0 seconds to 2 seconds, the probes were given time to settle and therefore become stabilised, before the probe and target sequence hybrid melted.

The importance of melting the products within a specific range, dictated by the number of SNPs present was especially noticeable in the assay for ST-61 based on *unc*A17. In this assay a perfect match was seen at 65°C. However alleles other than *unc*A17 had up to three base changes within the area of the sensor probe. This meant that mismatched products melted at substantially lower temperatures than the perfect match. These mismatched sample peaks were missed in the early reactions due to the melt curve starting at 50°C. The assay was subsequently carried out with the starting temperature lowered to 45°C.

The failure of the Lightcycler reactions when using extracted genomic DNA rather than PCR products is difficult to explain. It could be attributable to the DNA supercoiling making it difficult for the primers and probes to bind to the target region, or due to the vast array of components and genomic DNA present within the reaction mix which made it difficult for the primer and probes to hybridise. Despite attempts to resolve this by heating of the DNA to release the supercoiling, purifing and serially diluting the samples, no satisfactory results were obtained.

There were a few other negative aspects to the strategy based on the Lightcycler. Primarily the low throughput was of concern. With the lightcycler capacity of 32 samples in one run and the SNP strategy requiring six separate PCR reactions for

identification of these five clonal complexes, only five samples (plus controls) could be analysed in one run. Although the lightcycler only took one hour to run, the additional time required for the first round PCR amplification step, would make the whole process time consuming. Additionally, the manual data analysis on the lightcycler by reading the melting peak heights was also a time consuming step. Another consideration was the time consuming optimisation required for each probe and primer combination. If this strategy were to be extended to incorporate additional predictive alleles and clonal complexes then a great deal of time would be required to achieve optimal results.

To date, no studies have been carried out with respect to the relationship and distribution of SNPs within the MLST alleles of the *C. jejuni* scheme, however a recent study in Australia (Robertson *et al.*, 2004) interrogated the relationship of SNPs within the *N. meningitidis* and *S. aureus* MLST schemes. This study was a bioinformatics driven investigation whereby a specialist computer program was used to define highly informative sets of SNPs within the MLST databases. The authors described the necessity for the detection of between two and fourteen SNPs per ST to be identified, and the use of the Lightcycler real time platform utilising SYBR green chemistry for this detection. One limitation of using the Lightcycler with SYBR Green is the inability to multiplex, therefore for detection of each SNP separate reactions would be required. This would seem acceptable for the detection of two SNPs to define a ST, however the use of 14 separate Lightcycler reactions (described in next section) for the identification of one ST would be impractical, costly and unsuitable for rapid detection strategies. In addition, the relationships described for the *N.meningitidis* SNPs and STs are only applicable to the Australian MLST

database, similar relationships may not be defined within the UK based *N.meningitidis* MLST database.

The methods described have been shown to be specific and sensitive for three target MLST clonal complexes and the inclusion of additional alleles to clonal complexes ST-45 and ST-257 would provide equal specificity. However a major downside of the technique is that the SNP assays cannot be performed directly on isolated genomic DNA. The nested approach whereby the PCR product is amplified before the SNP Lightcycler PCR is carried out may aid the specificity of the assays, however in the context of the implementation of a rapid screening strategy for epidemiological studies, this is not ideal. A preferable strategy would be to have a SNP characterisation strategy, which was highly specific directly from genomic DNA.

Several observations concerning the performance of the SNP assays on the Lightcycler and practical aspects relating to the machine meant that the lightcyler was not considered the most suitable platform on which to base these assays. The new Lightcycler 2 instrument offers potential advantages over the existing system, the most promising feature being its capability of multiplexing with four detection channels and automatic melting peak calling. However it still remains a low throughput instrument with only 32 samples per run. Therefore the strategy was not developed further based upon this platform and the Taqman system was considered for further development. With the capability of SNP detection through the use of Minor Groove Binding probes, which offer, increased specificity (www.appliedbiosystems.com), and the possibility of running 96 samples in one run,

these assays on the Taqman would provide the desired high throughput, which is described in the next chapter.

Table 7.4 Example of results of SNP assays for clonal complex ST-21 for the MLST reference strain collection, identification of ST-21 assigned isolate by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP T_m Results (°C)			ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	Clonal complex	*glnA*1 108	*glnA*1 267	*tkt*_1 330	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	68	68	66	21
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	66	68	64	NA
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	66	68	66	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	62	64	64	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	62	64	66	NA
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	66	68	60	NA
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	66	68	64	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	66	68	62	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	62	68	64	NA
13265	human stool	8	10	2	2	11	12	6	354	ST-354 complex	66	68	64	NA
13255	human stool	1	3	6	4	3	3	3	22	ST-22 complex	62	64	64	NA
13266	human stool	1	2	49	4	11	66	8	362	ST-362 complex	66	68	64	NA
13259	human stool	3	1	5	17	11	11	6	49	ST-49 complex	68	68	64	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	66	64	64	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	62	68	62	NA

***Shaded boxes indicating highest melting temperatures for each SNP**

Table 7.7 Example of results of SNP assays for clonal complex ST-45 for the MLST reference strain collection, identification of ST-45 assigned isolate by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP T_m Results (°C)		ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	Clonal complex	*glt*A10 201	*glt*A10 225	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	61	65	NA
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	66	65	45
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	61	65	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	61	64	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	61	65	NA
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	57	64	NA
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	60	65	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	66	64	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	61	64	NA
13265	human stool	8	10	2	2	11	12	6	354	ST-354 complex	61	64	NA
13255	human stool	1	3	6	4	3	3	3	22	ST-22 complex	61	65	NA
13266	human stool	1	2	49	4	11	66	8	362	ST-362 complex	61	65	NA
13259	human stool	3	1	5	17	11	11	6	49	ST-49 complex	61	65	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	61	64	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	61	65	NA

***Shaded boxes indicating highest melting temperatures for each SNP**

Table 7.10 Example of results of SNP assays for ST-48 and 61 for the MLST Reference Strain Collection, identification of ST-48 and 61 assigned isolates by MLST and SNP analyses.

Isolate	Source	Allelic Profile								Clonal complex	SNP T_m Results (°C)				ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST		*glnA*4 18	*glnA*4 202	*unc*A5 186,189	*unc*A17 336	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	58	60	65	50	NA
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	61	60	63	50	NA
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	65	60	65	50	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	65	65	63	65	61
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	65	65	65	50	48
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	58	60	63	50	NA
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	58	60	63	50	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	58	60	63	50	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	58	60	63	50	NA
13265	human stool	8	10	2	2	11	12	6	354	ST-354 complex	58	65	63	50	NA
13255	human stool	1	3	6	4	3	3	3	22	ST-22 complex	61	65	63	50	NA
13266	human stool	1	2	49	4	11	66	8	362	ST-362 complex	58	60	63	50	NA
13259	human stool	3	1	5	17	11	11	6	49	ST-49 complex	59	60	63	50	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	65	60	65	50	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	58	60	63	50	NA

***Shaded boxes indicating highest melting temperatures for each SNP**

Table 7.13 Example of results of SNP assays for clonal complex ST-257 for the MLST reference strain collection, identification of ST-257 assigned isolate by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP T_m Results (°C)	ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	ST-Complex	*glyA*62 483	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	61	NA
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	61	NA
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	61	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	61	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	61	NA
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	66	257
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	61	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	61	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	61	NA
13265	human stool	8	10	2	2	11	12	6	354	ST-354 complex	61	NA
13255	human stool	1	3	6	4	3	3	3	22	ST-22 complex	61	NA
13266	human stool	1	2	49	4	11	66	8	362	ST-362 complex	61	NA
13259	human stool	3	1	5	17	11	11	6	49	ST-49 complex	61	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	61	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	61	NA

***Shaded boxes indicating highest melting temperatures for each SNP**

Table 7.16 Summary of results for 221 isolates showing the highest melting temperatures for each assay and indicative of the clonal complex. ND-not determined by this assay. [1] Clonal complex determined by full MLST (Dingle *et al* 2001). [2]Clonal complex determined by Lightcycler melting peak profiles *Any combination of metling temperatures possible. (Red indicates discrepancies in SNP assay results and full MLST. Shaded areas indicating highest melting temperatures indicative of the clonal complex)

C. jejuni MLST clonal complex[1] (number of isolates)	Melting Peak Temperatures for each SNP (°C)										Designated Clonal Complex[2]
	ST-21			ST-45		ST-257	ST-61 & 48				
	glnA1 108	*glnA1* 267	*tkt_1* 330	*gltA10* 201	*gltA10* 225	*glyA62* 483	*glnA4* 18	*glnA4* 202	*unc*A5 186,189	*unc*A17 336	
ST-21 (51)	68	68	66	66/61	65	66/65/61	60/58	65/58	65	50	**ST-21 (51)**
ST-257 (29)	66	68	60	57	64	66	60/58	60/58	63	50	**ST-257 (33)**
ST-61 (6)	61	64	66/64	61	64	61	65	65	63	65	**ST-61 (6)**
ST-48 (21)	62/61	64/68	66/62/60	61	65	61	65	65	65	50	**ST-48 (21)**
ST-45 (28)	66	68	66/64/60	66	65	66/65/61	62/60	61/58	63	50	**ST-45 (30)**
ST-206 (5)	66/61	68/64	66	61	65	61	65/60	65/58	65	50	ND
ST-49 (4)	68	68	64/62	61	65	61	60/59	60/59	63	50	ND
ST-22 (7)	62/61	64	66/64	61	65	61	65/61/60	65/62/60	63	50	ND
ST-177 (2)	66	68	64/62	66	64/61	61	60/58	60	63	50	ND
ST-353 (4)	66/62	68	65/62	61	65	61	60	61	63	50	ND
ST-42 (10)	66	68	64/62/60	60	65	61	50	60/58	63	50	ND
ST-52 (7)	62/61	68	66/65/64	61	64	61	50	60/58	63	50	ND
ST-354 (10)	66/62	68	66/64/62	61/62	65/64/61	61	65/60/58	65/61/58	63/62	50	ND
ST-403 (4)	66	64	64/62	62/61	65/64	61	65/61	65/60	65/63	50	ND
ST-362 (1)	62	68	62	61	65	61	58	60	63	50	ND
Other and unassigned (32)	*	*	*	*	*	*	*	*	*	*	ND

Chapter 8

Design of assays for strain specific detection and characterisation based on the Taqman.

Chapter 8

Design of assays for strain specific detection and characterisation based on the Taqman

8.1 Introduction

Various technologies are available for rapid sequence analysis and the detection of SNPs, essential for a strategy aimed at fast detection and characterisation of *C. jejuni*. The limitations of the Lightcycler real time PCR platform meant that the SNP assays were redesigned and retested for another SNP detection platform to achieve greater throughput and accuracy.

The Taqman is an alternative real time PCR system, which has the capability of accurate SNP detection through the use of specially designed probes known as Minor Groove Binding (MGB) probes (Applied Biosystems 2002). Studies have been described where the Taqman system has been employed for the rapid detection of SNPs. These have been predominantly within human genetics for large scale screening studies of genetic diseases for example (Jurevic *et al.*, 2003; McGuigan & Ralston 2002; Ranade *et al.*, 2001; Ward *et al.*, 2000).

SNP detection on the Taqman is based upon 5'→3' exonuclease activity of *Taq* polymerase (Holland *et al.*, 1991). Two fluorogenic probes can be used with a 5' reporter dye (FAM or VIC) and a 3' quencher. The probe is designed to be complementary to the sequence of interest, and located directly over the area of the SNP. When the probe is intact the fluorescence of the reporter dye is suppressed due to its close proximity to the quencher. However during the PCR reaction the reporter dye is cleaved by the action of the *Taq* polymerase resulting in an increase in

fluorescence, which is detected by the system. A perfect match between the probe and the target sequence allows successful hybridisation of the probe to the target strand, allowing the *Taq* to cleave the reporter dye and resulting in an increase in fluorescence signal. Any mismatches between the probe and target sequence reduce the efficiency of the probe hybridisation and subsequent cleavage cannot occur effectively. As a result, there is a substantial decrease in the amount of fluorescent emissions, which are barely detectable (if at all) by the system.

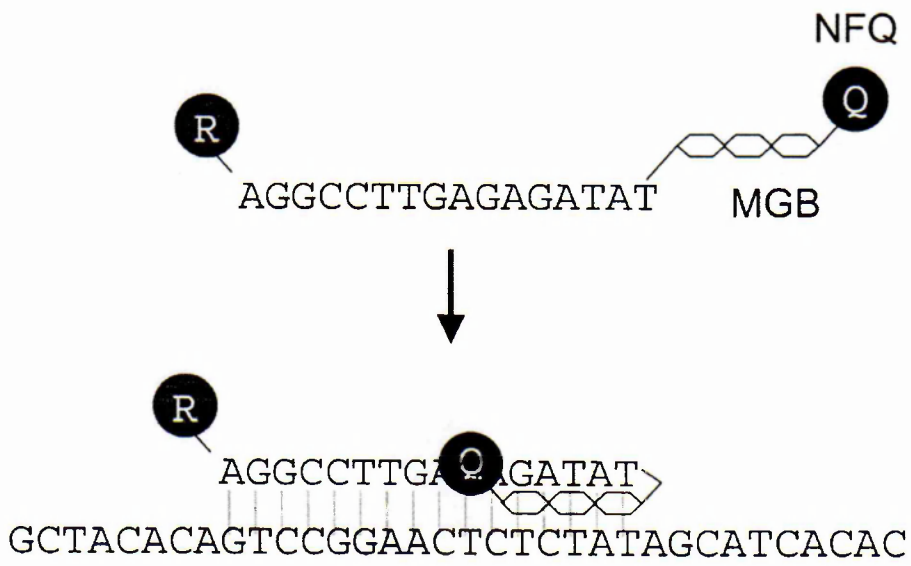
In many Taqman SNP assays described to date, biallelic SNPs are identified for confirmation of two specific genotypes. Two Taqman probes would be used in a bi-allelic system, each probe being specific to an allelic variant and labelled with a different reporter dye (FAM or VIC). After amplification the allelic content of the sample would be determined by the relative contribution of each reporter dye. However for the purposes of these assays, it was not necessary to determine both designations of a biallelic SNP, only the SNP of interest to identify an allele, therefore only one probe was used per SNP to be detected.

8.1.1 Minor Groove Binding (MGB) probes

The Taqman Minor Groove Binding (MGB) chemistry was used for these assays, which offered potential advantages for Taqman assays designed for SNP detection. MGB probes were similar to the standard Taqman probes (described in Chapter 4), however they also possessed an exclusive MGB component to dramatically increase the stability and specificity of probe hybridisation. Additionally they also used a non-fluorescent quencher in place of the previous standard TAMRA quencher, which improved the spectral performance (Applied Biosystems 2002).

The MGB group forms an extremely stable duplex with single stranded DNA targets allowing shorter probes to be used in the hybridisation assays (Figure 8.1). The MGB is a small crescent shaped molecule attached at the 3' end along with the quencher. When the Taqman probe hybridises, the MGB group fits into the minor groove of duplex DNA created between the probe and the target sequence, and is stabilised, mainly by van der Waals forces (Holland *et al.*, 1991). Stabilisation was highly effective when the sequences are perfectly matched. The increased stability allows the Taqman probes to be much shorter, around 13-20mer compared to the longer Taqman probes (20-24mer) described in chapter 4.

Fig 8.1 Taqman MGB probe and on hybridisation with target sequence (Applied Biosystems 2002) **R**- reporter dye (FAM or VIC), **Q**- quencher dye, **MGB**- minor groove binding component, **NFQ**- non fluorescent quencher.



The non-fluorescent quencher (NFQ) dye (or dark quencher) is a chromophore, which acts as an energy transfer acceptor from the reporter molecule (FAM or VIC), and did not emit a detectable fluorescent signal of its own. This has the advantage of producing a less complicated signal with lower fluorescent background, improving spectral discrimination and making the data interpretation easier (Applied Biosystems 2002).

This chapter explores the development of SNP assays based on the Taqman system using redesigned primers and MGB probes for the Taqman platform. Using this platform for the SNP based assays had many potential benefits over the strategy based on the Lightcycler. For these reasons the strategy described here based on the

Taqman was more comprehensive. The strategy was extended to include the clonal complex ST-206 and additional predictive alleles and associated SNPs were incorporated to improve the specificity and accuracy. The assays were also validated with the MLST reference strains (Wareing *et al.*, 2003) and the collection of diverse human animal, and environmental isolates as described in the previous chapter. Moreover the SNP strategy was tested for its use as a detection method for specific *C. jejuni* strains. This included testing the assays for specificity on a range of different *Campylobacter* species and related organisms and the testing of potential sources of infection including naturally contaminated meat and water samples.

8.2 Materials and Methods

8.2.1 Isolates selected and sample types tested

Cultured Samples

- NCTC *Campylobacter* species (National Collection of Type Cultures, London UK.) were used for assessing the specificity of assays.
- MLST reference strains (Set 1-section 7.2.1) (Wareing *et al.*, 2003)
- *C. jejuni* MLST characterised strains from varied sources (Set 2 -section 7.2.1).

Samples for direct detection

- *C. jejuni* confirmed naturally contaminated meat samples (direct DNA extraction by MagNApure) (Section 4.3.11).
- *C. jejuni* confirmed naturally contaminated water samples (direct DNA extraction by Prepman Extraction Reagent) (Section 2.2)

Methods

All SNP assays designed on the Taqman for six clonal complexes were tested for specificity for *C. jejuni* by using NCTC strains of other members of the genus campylobacteria, related organisms and other enteric pathogens. Additionally all assays were tested for specificity for each clonal complex by use of the MLST reference strains as described by Wareing *et al* (2003), and tested on the larger sample set 2 as described for the Lightcycler assays (Chapter 7).

The SNP assays were also tested for their potential use as a detection method for specific *C. jejuni* strains by clonal complex. Naturally contaminated meat samples

and water samples were used for direct detection studies. DNA was isolated from the meat samples directly by use of the MagNApure (Chapter 4) and the presence of *C. jejuni* was confirmed by the Taqman speciation assay (Chapter 4). All meat samples, which were *C. jejuni* positive, were tested for their clonal complex, by the SNP strategy. To provide a confirmatory test that the SNP assays had correctly assigned the clonal complex, the *C. jejuni* positive samples, which were also culture confirmed *C. jejuni* were subjected to full MLST from culture. Some of the Taqman meat samples could not be culture confirmed, in these cases confirmatory testing was carried out by means of full MLST from the direct DNA extract.

Water samples taken during 2003 from rivers and lakes, had been tested for campylobacters by Dr. S. Lai from the Water and Environmental Reference Unit (HPA, Colindale). Water samples had been enriched in campylobacter enrichment broth and the Prepman DNA extraction reagent was used to isolate the DNA. Campylobacter presence had been confirmed by the use of the Taqman PCR assay as described in chapter 4 (Best *et al.*, 2003) by Dr S. Lai.

All *C. jejuni* positive samples were subsequently tested for their clonal complex by the SNP strategy using the direct DNA extracts. No culture was available; therefore confirmatory testing was carried out by full MLST on the same extract.

8.2.2 Design of Taqman primers and probes

Materials

- Primer Express version 1.5 (Applied Biosystems)
- Sequences of MLST alleles and SNP positions (identified in chapters 5 and 6)

Method

Assays were designed in either a uniplex or duplex format dependent on the number of SNPs to be detected and were similar to those described for the Lightcycler. The Taqman assays were less complicated to design due to there being only one probe and two primers required for detection of each SNP. Additionally the Taqman primers and probes were designed to work with standardised reaction conditions and PCR cycling, therefore no titration of $MgCl_2$ was required or extensive optimisation of cycling parameters. Primers and probes were designed using Primer Express as described in section 2, the specific criteria for successful amplification of MGB probes (described below), were included into the design parameters of the program.

Criteria for design of MGB probes (Applied Biosystems)

- No guanine residue at the 5'end of the probe
- An estimated T_m of 65-67°C
- As short as possible without being shorter than 13 nucleotides
- The polymorphic site should be positioned in the central third or towards the 3' end of the probe.

All primers and probes were tested in the Primer Express Test Document to ensure they had the correct melting temperatures. All probes were synthesized by Applied

Biosystems, and stored at -30°C until use. When ready to use, primers (MWG Biotech) and probes were diluted in sterile 1x TE buffer.

8.2.3 Allelic discrimination assays

Materials

- Sequence Detection (Taqman) System 7000 (Applied Biosystems)
- Taqman universal mastermix (Eurogentec)
- 5µM stock MGB probes (Applied Biosytems) (Table 8.1)
- 10µM stock Primers (MWG Biotech) (Table 8.1)
- Nuclease Free water
- Taqman 96 well optical reaction plates and adhesive covers (Applied Biosystems)
- *Campylobacter jejuni* genomic DNA (extracted by MagNApure)

Method

Taqman reactions were applied to MagNApure extracted DNA. All reactions were initially set up following manufacturers guidelines for maximum reaction efficiency and subsequently optimised for each specific assay. The newer model Taqman (version 7000) was used. This system superseded the original Taqman 7700 due to running on Windows platform with the potential to be integrated into laboratory networks and redesigned software. Each 25µl PCR reaction comprised 2.5µl MagNApure extracted DNA, 300µM forward and reverse primers, 100µM probe and 1x Taqman universal mastermix all added into Taqman 96 well reaction plates and sealed with adhesive covers.

The Taqman mastermix and 96 well Taqman reaction plates for SNP analysis were either set up manually by the use of multichannel pipettes, or by the use of a robotic liquid handling machine (Corbett Robotic Instrument CAS 1200). This robot was used for the preparation of mastermix, distribution of this into the 96 well Taqman Reaction plate and addition of DNA samples. The increased precision and accuracy of the machine meant that the total reaction volumes could be reduced to a quarter of the recommended reaction volumes without compromising quality. In these cases a total reaction volume of 12.5µl was used incorporating 1.25µl DNA sample.

Universal cycling conditions were used for most assays however the annealing temperature and time was adjusted for some of the assays to achieve better discrimination. Cycling comprised 10 minutes at 50°C to activate the enzyme UNG, 10 minutes at 95 to activate the *Taq* polymerase followed by 40 cycles of denaturation at 95°C for 1 minute and annealing/extension at 65°C for 1-1.20 minutes.

8.2.4 Data analysis

Data analysis was carried out as described in Chapter 4, the presence of the informative SNP was recognisable with a C_T of between 14-20. For every assay optimal reaction conditions were determined by the effect on C_T by changing the reaction conditions, in particular the annealing/extension time and temperature, which improved the discrimination in some assays. It was attempted to maintain the standardised annealing temperatures for each assay at 65°C for 1.20 minutes to allow for all the assays to ultimately be run on one plate, allowing for a panel of reactions for all the complexes to be created.

8.3 Results and Discussion

8.3.1 General assay design

Primers and probes were designed to detect the SNPs using similar assay structure to the strategy on the Lightcycler. The main differences were that additional alleles were incorporated to improve the specificity for ST-45 and ST-257 clonal complexes and the strategy was extended to include the ST-206 clonal complex. In total, eight assays were designed to detect the six target clonal complexes. These comprised a duplex and uniplex assay for ST-21, a duplex and uniplex for ST-45, two duplex reactions for ST-48 and ST-61, a uniplex reaction for ST-206 and a duplex reaction for ST-257. The Taqman was based upon different chemistries to the Lightcycler, therefore there was no requirement to position the primers in conserved areas. This meant that placing primers within divergent regions of the predictive allele increased the overall specificity of the assay for that particular allele. Each assay was optimised by variation in the annealing/extension temperature within the range of 60-67°C and also by increasing the time from 1 minute to 1 minute and twenty seconds. Probe and primer sequences are shown in table 8.1.

Positive results indicating the presence of the SNP were confirmed by a C_T result within the range of 13-20, negative results were indicated as a C_T value of 40. For some samples C_T values within the range of 30-40 were also seen, but confirmed as negative. For the direct detection studies C_T values above these were seen due to the smaller amount of DNA present within non-cultured samples. In these cases the values considered positive were up to the value of 38 as described for the Taqman assay in Chapter 4.

Table 8.1 Taqman primer and probe sequences for SNP clonal complex assays R –reverse primer, F- forward primer, FAM reporter dye, VIC reporter dye

Assay	Alleles and position of SNP(s)	Type of Assay	Primer sequences	Probe sequences
ST-21	*gln*A1 108, 267	duplex	**ST21*gln*108F** AAATGTCCAAGAAGCATAGCAAAA **ST21*gln*108R** AAG TAAGCAGTATCAGCTATGCCACTA **ST21*gln*267R** CCGAAGAAGGAGAGTGGAATGA **ST21*gln*267R** GAACTGGAAAATATCCACCTTTGTTT	**ST21*gln*108FAM** FAM-AAA GCA ATA GAA CAC TT **ST21*gln*267VIC** VIC-AAT TTA CCG ATA GCT ACA ATA
	*tkt*_1 330	uniplex	**ST21*tkt*330F** GGACTTCACAAACTTGATAACTTCATACTT **ST21*tkt*330R** CATTTTCGTTAAAGGCTAAACCTACA	**ST21*tkt*330FAM** FAM-CAA TAT CTC TAT AGA AGG CGA
ST-45	*glt*A10 201, 225	duplex	**ST45*glt*201F** AATATATGGAAATGGCAGCTAGAATAGTAG **ST45*glt*201R** CGATCTAAATTTGGATAAGCCATAGG	**ST45*glt*201FAM** FAM-ACT ATA GTC GCC ACC GC **ST45*glt*225VIC** VIC-AGC CGT GTT TAT ATC T
	tkt_7 138, 141	uniplex	**ST45*tkt*138F** GCTACAGGCCCTTTAGGACAAG **ST45*tkt*138R** TTTTTGCTGCCATAGCAAAGC	**ST45*tkt*138FAM** FAM-CGT TGC AAA CGC TGT A

ST-257	***gly*A62 483**	duplex	**ST257*gly*62F** TGATGCATGTAATTGCTGCAAA **ST257*gly*62R** AAAACTTGAGCATTGGTTCTTACTTG	**ST257*gly*62VIC** VIC -TTTGCATAGTAGACTTTCCAC
	***pgm*_4 165**		**ST257*pgm*163F** GGCGTATTGGCACTTTATCTTAAAG **ST257*pgm*163R** CGCCAACATTACAAGTATCAAGTTCT	**ST257*pgm*163VIC** FAM-AAG TGT TGT GGC TAC TAT AA
ST-61	***gln*A4 18, 202**	duplex	***gln*4-18F** AAAACCAGCTGATGCTCAAAGTGCTT ***gln*4-201R** CCACTCTCCTTCTTCGGTATCAA	***gln*4-18FAM** FAM-CAG CTG ACC CTA CTA T ***gln*4-220VIC** VIC-ATA CTG CTC ATT GTT C
	***unc*A17 336**	duplex with *uncA5*	**ST61*unc*336F** TTTCAGCTTATATTCCAACCAATGTT **ST61*unc*336R** TGGACGAATTCCTGAGTTAAATAAATC	**ST61*unc*VIC** VIC-ATT TCG ATC ACT GAT GGA
ST-48	***gln*A4 18, 202**	duplex	As above	
	***unc*A5 186,189**	duplex with *uncA17*	**ST48*unc*189F** TCCTCCAGGTCGTGAAGCTT **ST48*unc*189R** AGAACCAGCACCTAATTCATCATTT	**ST48*unc*5FAM** FAM-CAG GTG ATG TTT TTT ACC TT
ST-206	***tkt_1* 330**	uniplex	As above	
	***glnA21* 18, 33**	uniplex	**ST206*gln*18F** AAAACCTGATGCTCAAAGTGCTT **ST206*gln*18R** TTTTGCTATGCTTCTTGGACATTT	**ST206*gln*21FAM** VIC-ACC CTA CTA TCA TAG TAT TTT

8.3.2 Specificity Studies

NCTC strains of other members of the genus *Campylobacter* were tested as well as related organisms such as *Arcobacter* and *Helicobacter* and other enteric pathogens including *E. coli* and *S. typhi.* The SNP strategy had to be specific for *C. jejuni* to be suitable as a detection method and it was important that no other *Campylobacter* species were positive, and therefore detectable. All organisms tested other than *C. jejuni* were negative for each SNP assay (Table 8.2). It was important that no false positives occurred as other species of *Campylobacter* could be potentially isolated in the place of *C. jejuni.* Some species of *Helicobacter* may be potentially isolated due to having similar growth requirements and similarly *Arcobacter* are prevalent within chicken samples (Atabay & Corry, 1997). The NCTC *C. jejuni* strain (NC11351-01) was identified as belonging to clonal complex ST-21 by the SNP assays. This was also confirmed through full MLST.

Table 8.2 Specificity of the assay using DNA from other microrganisms and *Campylobacter* species (shaded areas indicating positives)

Genus	Species	NCTC/ref no.	Source	Mean Taqman result from all SNP assays (C_T)
Campylobacter species	*C. jejuni* ss *jejuni*	NC11351-01	NCTC	17.4*
	C. jejuni ss *doylei*	NC11847-04	NCTC	40
	C.coli	NC11353-04	NCTC	40
	C. fecalis	NC11415-04	NCTC	40
	C. lari	NC11352-07	NCTC	40
	C. consiscus	NC11485-06	NCTC	40
	C. fetus	NC10842-07	NCTC	40
	C. fetus	NC10354-07	NCTC	40
	G. hyointestinalis	NC11608-06	NCTC	40
	C. helveticus	NC12470-03	NCTC	40
	C. curvus	NC11649-03	NCTC	40
	C. sputorum ss *mucosalis*	NC11000-05	NCTC	40
	C. sputorum ss bubulus	NC11367-03	NCTC	40
	C. sputorum	NC11528-05	NCTC	40
	C. upsaliensis	NC11541-05	NCTC	40
Other organisms	*E. coli*	NC09001-19	NCTC	40
	S. typhi	ATCC-4543	NCTC	40
	A. butzleri	NC12481-03	NCTC	40
	A. cryaerophilus	NC11885-04	NCTC	40
	A. skirrowski	NC12713-02	NCTC	40
	A.nitrofigilis	NC12251-02	NCTC	40
	H. pylori	NC11637	LEP	40
	H. pullorum	C77339	LEP	40
	H. canadensis	C78661	LEP	40

***mean C_T result calculated for three informative SNPs for the ST-21 clonal complex. Results for all other assays were negative. Negative result-C_T value of 40. Positive result C_T within range of 13-20.**

8.3.3 SNP assay for clonal complex ST-21

(i) Design of primers and probes

The alleles *gln*A1 and *tkt*_1 had been identified to be suitable representative alleles for this clonal complex. The associated SNPs are summarised below.

Allele	SNP
glnA1	**A→ G** (108bp)
	C→T/A (267bp)
tkt_1	**T→C/G** (330bp)

Design of the assay for clonal complex ST-21 was very similar to the assay design on the Lightcycler. Two assays were designed, (i) a duplex assay for the detection of the two SNPs on the *gln*A1 allele and (ii) a separate assay for the SNP on the *tkt*_1 allele. For the duplex *gln*A1 assay one set of forward (75-100bp) and reverse (121-149bp) primers were designed and also an internal MGB probe labelled with FAM to detect the SNP at 108bp. For the SNP at 267bp, forward (230-250bp) and reverse (300-322bp) primers were designed and a MGB probe labelled with the fluorescent dye VIC. (Table 8.1)

For the uniplex *tkt*_1 assay to detect the SNP at 330bp, forward (276-306bp) and reverse (341-368bp) primers were designed and an internal MGB probe labelled with FAM.

(ii) Results for reference isolates

Presence of the SNP on *gln*A1 at 108bp was confirmed by a C_T value of 15.5-16, which was present within isolates 13254 (ST-21), and 13259 (ST-49). Throughout the remainder of the reference isolates the results were negative (C_T of 40) indicating the presence of a base change, therefore no binding of the probe and no detectable

increase in fluorescence. At this position (108bp) the C_T value of 16-17 was due to the presence of an A at 108bp, with the C_T value of 40 corresponding to the presence of a G at this position (Table 8.3).

The SNP at *gln*A1 267bp was more frequent within the alleles of the reference set and a low C_T of 16-17 was seen for more than half of the reference isolates including isolates 13254 (ST-21) and 13259 (ST-49) (Table 8.3). The low C_T was due to the presence of a C at 267bp. The negative C_T values (40) within isolates 13261 (ST-61), 13258 (ST-48), 13255 (ST-22) and 13268 (ST-403) were due to these alleles containing a T at position 267bp. No samples in this set contained alleles with the alternative base change to A at this position, if any of these alleles had been present it would be likely that a C_T of 40 would also be obtained due to the extra sequence divergence around this base change, resulting in minimal probe binding.

The final SNP in the strategy for identification of the clonal complex ST-21 was *tkt*_1 330bp, which was also seen within the reference isolates including ST-21 but not including ST-49. This SNP was crucial for the distinction of ST-21, as both clonal complexes shared the initial allele *gln*A1. The low C_T values represented the presence of the T at 330bp; a C_T of 40 represented the presence of a C at 330bp. Within this sample set no other base change of the triallelic SNP was present (Table 8.3).

As demonstrated by these results utilising the reference isolates for the *C. jejuni* MLST scheme it has been possible to identify the ST-21 clonal complex assigned isolate (13254) by the low C_T value obtained for each of these SNPs, which is only

seen in the ST-21 assigned isolate. From this data set no other relationships can be seen with respect to the designation of other isolates into clonal complexes.

(iii) Results for application of the SNP assay for clonal complex ST-21 to larger sample numbers

The assay was still valid by application to a larger sample set encompassing strains from diverse sources. All the ST-21 clonal complex isolates confirmed by full MLST were correctly identified by the SNP assay (Table 8.4). The same low C_T profile for the three SNPs was consistent across the ST-21 clonal complex isolates with little variation in C_T values for positive samples. Mean C_T values calculated for each SNP ranged from 16.1-18.4.

The large and diverse nature of this clonal complex necessitated the use of two predictive alleles for accurate identification, which was successfully demonstrated on the Lightcycler and on this platform. It was considered that it may be necessary to increase the number of loci used to improve the resolution, however for the dataset studied the specificity was adequate and all the ST-21 clonal complex isolates were successfully identified by the two predictive alleles. With all the SNP assays it could be possible to increase specificity by the incorporation of additional alleles, however, there comes a point where little additional information is attained for the expense and effort involved.

Table 8.4 Results for all ST-21 assigned isolates (n=51) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP result (C_T value)		
										*gln*A1 108	*gln*A1 267	*tkt*_1 330
47762	chicken offal or meat	4	1	1	3	1	1	5	712	15.8	15.2	17.6
47771	chicken offal or meat	2	1	1	3	2	1	5	21	16.64	16.8	19.2
47785	chicken offal or meat	4	1	1	3	4	1	5	107	16.8	16.4	17.6
47860	human stool	2	1	21	3	2	1	5	53	14.7	14.7	20.4
48237	human stool	2	1	1	3	2	1	5	21	16.1	18.3	17.1
48279	human stool	2	1	21	3	2	1	5	53	14.1	15.1	18.7
48398	chicken offal or meat	2	1	5	3	2	1	5	19	16.7	16.3	17.7
48474	chicken offal or meat	2	1	5	3	2	1	5	19	15.8	15.3	17.6
48479	chicken offal or meat	2	1	6	3	4	1	1	722	17	16	17.7
48540	human stool	2	1	21	3	2	1	5	53	15.1	15	17.7
48710	human stool	2	1	21	3	2	1	5	53	15	19	20
48825	human stool	2	1	1	3	7	1	5	104	16.5	18.3	19
48832	human stool	2	1	1	3	7	1	5	104	15	15.1	18.1
49006	human stool	2	1	21	3	2	1	5	53	14.9	15	18.6
49390	human stool	2	1	1	3	2	1	5	21	15	16	18
49428	human stool	2	1	1	3	2	1	5	21	16.1	18	17
49438	human stool	2	1	1	3	2	1	5	21	16.2	17	19.2
49546	human stool	2	1	21	3	2	1	5	53	16.5	19.6	20
49600	human stool	2	1	5	3	2	1	5	19	18.4	20	19.5
49688	human stool	2	1	5	3	2	1	5	19	16.4	16	16.7
49828	human stool	2	1	1	3	2	1	5	21	16	15.2	18.4
49938	human stool	2	1	5	3	2	1	5	19	15.2	15.1	18.8
50258	chicken offal or meat	8	1	21	3	2	1	5	119	16	18	18.7
50314	human stool	2	1	1	3	2	1	5	21	16.1	18	18.9
50707	human stool	2	1	1	3	7	1	5	104	14.4	14.5	19
50785	human stool	8	1	6	3	2	1	1	44	16.3	16.2	17
50815	chicken offal or meat	2	1	5	3	2	3	5	190	20	20	17.2
50963	human stool	2	1	1	3	2	1	5	21	16.3	18.1	15.8
51431	chicken offal or meat	9	1	21	3	2	1	5	347	16.1	15.3	16.5
51689	Human stool	2	1	1	3	2	1	5	21	16.3	17.9	17.7
51763	chicken offal or meat	2	1	1	3	7	1	5	104	15.7	16.3	20
51820	chicken offal or meat	8	1	6	3	2	1	1	44	15.8	16.3	18.7
52155	human stool	2	1	1	4	2	1	5	761	16	15.1	17.8
52376	human stool	2	1	1	3	2	1	5	21	16.8	15.2	16
52498	chicken offal or meat	7	1	1	3	2	1	5	615	16.1	18	19.5
52530	chicken offal or meat	8	1	6	3	2	1	1	44	17.5	14	20
52599	chicken offal or meat	7	1	1	9	7	1	5	742	16	17	16.2
52724	chicken offal or meat	4	1	1	3	7	1	5	744	15.9	15.5	16.4
52833	human stool	2	1	1	3	2	1	5	21	15.8	15	19.7
53211	human stool	2	1	12	3	2	1	5	50	15.1	18	17.4
53532	human stool	2	1	1	3	2	1	5	21	16	14.3	16.2
53793	chicken offal or meat	4	1	1	3	2	1	5	640	17.1	20	17.6
53798	chicken offal or meat	8	1	1	62	7	1	5	748	16.7	14.2	19.4

54068	chicken offal or meat	8	1	5	4	2	1	5	752	14.1	14.3	16.4
54349	human stool	2	1	1	3	2	1	5	21	16	19.2	18.7
54743	human stool	2	1	1	3	2	1	5	21	15.5	14	20.1
54984	human stool	2	1	1	3	2	1	5	21	16.2	16	19.1
55053	human stool	2	1	21	3	2	1	5	53	17.6	16.3	17.6
55142	human stool	2	1	21	3	2	1	5	53	16.4	16.1	17.6
11168	human	2	1	5	3	4	1	5	43	15.8	15.3	17.5
13254	beef offal	2	1	1	3	2	1	5	21	18.2	17.1	17.2
Mean C_T values										16.1	16.4	18.4

8.3.4 SNP assay for clonal complex ST-45

(i) Design of primers and probes for clonal complex ST-45

The alleles *glt*A10 and *tkt*_7 had been identified as representative alleles for this clonal complex. The associated SNPs are listed below.

Allele	SNPs
***glt*A10**	C→G/A (201bp)
	C→T (225bp)
***tkt*_7**	A→C (138bp)
	C→T (141bp)

The design of this assay on the Taqman differed to the assay design on the Lightcycler by the addition of the extra two SNPs on the *tkt*_7 allele to improve the specificity of the assay, and to reduce the chance of detecting false positives due to the reliance on just one allele. Two assays were designed on the Taqman. A duplex assay to detect the two SNPs on the *glt*A10 allele was designed similar to the Lightcycler assay and an additional uniplex assay to detect the SNPs on the *tkt*_7 allele. The close proximity of the two SNPs on the *tkt*_7 allele made it unnecessary to use two probes as both could be detected by the use of one probe. For the duplex *glt*A10 assay one set of forward (150-181bp) and reverse (231-257bp) primers was designed and two internal MGB probes labelled with FAM to detect the SNP at 201bp and VIC to detect the SNP at 225bp. The same problems were encountered as in the case of the Lightcycler design due to the close proximity of the two SNPs. Initially it was considered that it was going to be impossible to fit two Taqman probes into such a small amount of sequence. Also no reports had been made about the success of using two Taqman probes in close proximity. The only way it could be designed was by the probes being positioned back to back (4bp gap) as in the

Lightcycler case and with shared forward and reverse primers, something that at the time of writing had not been described before (Table 8.1).

For the uniplex *tkt*_7 assay to detect the SNPs at 138bp and 141bp forward (109-131bp) and reverse (169-190bp) primers were designed and an internal MGB probe labelled with FAM (Table 8.1).

(ii) Results for the reference isolates

Presence of the SNP at 201 was confirmed by a C_T value within the range of 15.4-18, which was seen in the isolate 13257 (ST-45) and one other isolate 13262 (ST-177) due to having the allele *glt*A8 which also contained this SNP (Table 8.5). This SNP was crucial for the successful identification of the allele *glt*A10 due to its low distribution throughout the remainder of the *glt*A alleles.

Presence of the SNP on *glt*A10 at 225bp was confirmed by a low C_T value of 15-16 as in the case of the ST-21 assay. This SNP was frequent throughout the majority of the *glt*A alleles within the reference set except for the isolate 13262 (ST-177) which had a *glt*A8 allele and displayed a base change at this position, resulting in unsuccessful probe binding.

Presence of the SNP on the *tkt*_7 allele was indicated by the low C_T value of 17.2-18 for isolates 13257 (ST-45) and 13268 (ST-403). In conjunction with the identification of the two SNPs on the *glt*A allele it was possible to successfully identify the clonal complex ST-45 isolate by the presence of this SNP.

(iii) Results for application of the SNP assay for clonal complex ST-45 to larger sample numbers

Identification of the ST-45 clonal complex was concordant with the full MLST results where 28 clonal complex ST-45 assigned isolates from the large sample set were successfully identified by the SNP assay. Mean low C_T values in the range of 15.7-17.9 were obtained for each of these SNPs and were indicative of this clonal complex (Table 8.6).

Improved specificity for this clonal complex was achieved by the addition of the further predictive allele *tkt*_7 into this strategy. This enabled only the ST-45 assigned isolated to be identified. In the Lightcycler strategy described in Chapter 7 two isolates were falsely identified as ST-45 by the presence of the *glt*A10 allele, when they were actually assigned into other clonal complexes. The additional predictive allele into the strategy prevented the inaccurate identification of the ST-45 clonal complex and also improved the validity of the assay by being based upon two different alleles.

Table 8.6 Results for all ST-45 assigned isolates (n=28) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP result (C_T value)		
										*glt*A10201	*glt*A10225	*tkt*_7138
47720	chicken offal or meat	24	7	10	62	42	7	1	756	15	14.5	20
47728	chicken offal or meat	2	7	10	4	1	7	1	233	15.8	15.1	17.1
47732	chicken offal or meat	2	7	10	4	1	7	1	233	15.1	14.5	16.4
47733	chicken offal or meat	9	7	10	4	1	7	1	320	14.9	14.4	15.3
48300	chicken offal or meat	33	7	10	4	1	7	1	714	15.1	15.1	17
48311	chicken offal or meat	4	7	10	4	1	7	1	45	14.7	14.2	17.4
48435	chicken offal or meat	2	7	10	4	1	7	1	233	15	14.6	18.1
48471	chicken offal or meat	2	2	10	4	2	7	1	721	16	18	17.9
48885	human stool	4	7	10	4	1	7	1	45	13	15	17
49294	human stool	4	60	10	4	1	7	1	765	16	18	16.2
49594	human stool	4	7	10	4	1	7	1	45	14.9	14.3	16.8
49979	human stool	4	17	10	4	1	7	1	766	16	18	17
50254	chicken offal or meat	4	4	10	4	2	7	1	734	18	19	17.1
50313	human stool	4	7	10	4	1	7	1	45	15	17	18.6
50725	human stool	4	7	10	4	1	7	1	45	14.7	14.2	18
51009	human stool	4	7	10	4	1	7	1	45	16	19	18
51405	chicken offal or meat	2	7	10	4	1	7	1	233	17	19	19.2
51953	human stool	4	7	10	4	1	7	5	241	14.6	14	20
53892	human stool	4	2	10	4	1	7	1	782	15	17	19.1
53967	chicken offal or meat	9	7	10	2	1	7	1	750	15	16.5	18.1
54401	human stool	4	7	10	1	1	7	1	25	19.9	18.8	18.7
54879	human stool	4	7	10	4	1	7	1	45	14.8	14.2	20.1
56050	chicken offal or meat	2	7	10	1	1	7	1	755	14.2	16	16.2
56052	chicken offal or meat	8	7	10	4	1	7	1	754	16	17.1	18.2
29615	wild bird	4	7	10	4	1	7	1	45	16	19.1	16.6
29645	Cattle	4	7	10	4	1	7	1	45	17	18.6	18.3
29703	canine	4	7	10	1	1	7	1	25	19.1	16.3	19.4
13257	human stool	4	7	10	4	1	7	1	45	16.2	15.3	19.6
Mean C_T values										15.7	16.3	17.9

8.3.5 SNP assays for clonal complexes ST-48 and ST-61

(i) Design of primers and probes

Identification of the clonal complexes ST-61 and ST-48 was with a combined strategy based initially on the shared predictive allele *gln*A4 then two specific predictive alleles for ST-48 and ST-61. This was similar to the strategy previously described for the lightcycler. The SNPs used are shown below.

Allele	SNPs
*gln*A4	C→T (18bp)
	G→A (202bp)
*unc*A17	C→T (336bp)
*unc*A5	T→C (189bp)

Two duplex assays were designed (i) a duplex assay for the detection of the two SNPs on the *gln*A4 allele and (ii) a second duplex assay for detection of the two SNPs on the *unc*A5 and *unc*A17 alleles. The two SNPs on the *gln*A4 allele (18bp and 202bp) were spaced apart therefore separate sets of primers were designed for each probe. For detection of the SNP at 18bp the allele sequence had to be extended by 50 bases as in the case of the lightcycler assay to fit in a primer before the location of the SNP. For this SNP forward (21-43bp) and reverse (111-140bp) primers were designed and an internal MGB probe labelled with FAM. For the SNP at 202bp forward (190-210bp) and reverse (272-295bp) primers were designed and internal MGB probe labelled with VIC (Table 8.1).

The second duplex assay to detect the SNPs at 333-336bp on the *unc*A17 allele comprised the design of forward (301-326bp) and reverse (364-390bp) primers and internal probe designed to include all the three SNPs and which was labelled with

FAM. This assay also included the design of forward (161-180bp) and reverse (236-261bp) primers for detection of the *unc*A5 allele and an internal MGB probe labelled with VIC (Table 8.1).

(ii) Results for the reference isolates

As in the case of the Lightcycler assay design for these clonal complexes, a combination of different positive C_T values was required for determination of each clonal complex these are shown below.

Clonal complex	Required range in C_T			
	*gln*A4		*unc*A5	*unc*A17
	18bp	201bp	190bp	336bp
For ST-61	<20	<20	40	<20
For ST-48	<20	<20	<20	40

Presence of the SNP on the *gln*A4 at 18bp was confirmed by a C_T value of 15.9-20, which was present within the isolates 13261 (ST-61) and 13258 (ST-48) of required clonal complexes and also within isolates 13263 (ST-206) and 13268 (ST-403). Throughout the remainder of the reference isolates the results were negative with a C_T of 40 indicating the presence of a base change therefore no binding of the probe and no detectable increase in fluorescence (Table 8.7).

Presence of the second SNP at 202bp on the *gln*A4 allele was also confirmed by a low C_T values, 17-19.4 seen in the isolates 13261 (ST-61) and 13258 (ST-48), however this SNP was also present in other isolates of the reference collection 13265

(ST-354) and 13255 (ST-22) due to the SNP being present in the alleles *gln*A10 and *gln*A3.

Presence of the SNPs on the *unc*A5 allele at 186 and 189bp were confirmed by a C_T value of 14.3-15, which was present within the isolates 13254 (ST-21), 13263 (ST-206), 13258 (ST-48) and 13268 (ST-403). The presence of the three SNPs (*gln*A4 18bp, *gln*A4 202bp and *unc*A5 186, 189bp) indicated the ST-48 clonal complex.

The specific SNPs required to detect the clonal complex ST-61 were at positions 333-336bp on the *unc*A17 allele, which was only seen in isolate 13261 (ST-61) from the reference collection. Due to these SNPs being completely unique for this allele this low C_T was not seen within any other isolates and also would be highly unlikely to occur outside of ST-61 within a larger strain set. This assay was also highly specific for the allele due to the location of the primers, which were also located on areas within this allele, which were highly diverse. This result in conjunction with the C_T obtained for *gln*A4 was able to confirm the assignment of the isolate 13261 into the ST-61 clonal complex. The presence of the three SNPs (*gln*A4 18bp, *gln*A4 201bp and *unc*A17 333bp) indicated the ST-61 clonal complex.

(iii) Results for application of the SNP assay for clonal complexes ST-48 and ST-61 to larger sample numbers

Using the combined approach, consisting of two duplex assays for ST-61 and ST-48 the complete set of ST-48 (n=28) and ST-61 (n=6) assigned isolates were correctly identified by the presence of specific SNPs from the dataset. For identification of clonal complex ST-48 both *gln*A4 SNPs and the *unc*A5 SNP, were positive and the

*unc*A17 SNP was negative for all isolates. Mean positive C_Ts within the range of 15.3-15.6 were obtained for the ST-48 SNPs, and a mean C_T of 40 was obtained for the ST-61 (*unc*A17) SNP (Table 8.8).

For identification of clonal complex ST-61 isolates both *gln*A4 SNPs and the *unc*A17 SNPs were positive and the *unc*A5 SNP was negative (40). Mean C_T values obtained for *gln*A4 allele amongst these isolates were higher than those obtained for the ST-48 assigned isolates (17.2 and 17.1) and a mean C_T of 17 was obtained for the *unc*A17 SNP. Additionally the *unc*A5 SNPs within these isolates was negative with a mean C_T of 40 (Table 8.9).

The combined approach for the identification of two clonal complexes was achievable due to the relatedness of the two clonal complexes by their sharing of predictive alleles. Therefore these assays based on the same predictive allele provided a more rapid and convenient system and reduced the number of reactions required in the overall clonal complex strategy, therefore reducing reagent costs. These assays on the Taqman when compared to the Lightcycler had taken relatively little set up time and optimisation. The two duplex reactions worked well with equimolar concentrations of primers and probes and with standardised cycling conditions.

Table 8.8 Results for all ST-48 assigned isolates (n=21) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP result (C_T value)			
										*gln*A4 18	*gln*A4 202	*unc*A5 191	*unc*A17 336
48538	human stool	2	4	1	2	7	51	5	205	16	15	15	40
49009	human stool	2	4	1	2	7	1	5	48	15.2	14.9	14.9	40
49474	human stool	2	4	1	2	7	1	5	48	15	14.8	14.8	40
50944	human stool	2	4	1	2	7	51	5	205	15	15	15	40
51531	human stool	2	4	1	2	7	51	5	205	15.7	15	15	40
51926	human stool	2	4	2	2	6	1	5	38	15.2	15.7	15.7	40
52142	human stool	2	4	1	2	7	1	5	48	15.4	15.2	16	40
52477	chicken offal or meat	9	4	1	2	7	1	5	738	15.9	16.1	15.2	40
52484	chicken offal or meat	9	4	1	2	7	1	5	738	15	17.1	17.1	40
52834	human stool	2	4	1	2	7	1	5	48	16	14	14	40
53086	chicken offal or meat	8	4	1	2	7	1	5	414	15.4	15.6	15.6	40
53703	chicken offal or meat	4	4	1	2	7	1	5	739	15.4	16.5	15.4	40
53706	chicken offal or meat	4	4	1	2	7	1	5	739	16.8	14.9	16.1	40
53785	chicken offal or meat	4	4	1	3	7	1	5	749	17.1	14.2	14.9	40
53810	chicken offal or meat	4	4	1	3	7	1	5	749	17.5	15.3	16.5	40
53854	human stool	2	4	1	2	7	51	5	205	15.6	17.2	15.2	40
54720	human stool	2	4	1	2	7	1	5	48	14.2	16.1	15	40
29697	ovine abortion	2	4	2	2	6	1	5	38	16	16.9	16	40
76792	human OB	2	4	5	2	7	1	5	66	17	15.8	16.8	40
76879	human OB	2	4	1	2	7	1	5	48	18	15.4	17	40
13258	lamb offal	2	4	1	2	7	1	5	48	19	16	16.4	40
Mean C_T values										15.3	15.5	15.6	40

Table 8.9 Results for all ST-61 assigned isolates (n=6) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP result (C_T value)			
										*gln*A4 18	*gln*A4 202	*unc*A5 191	*unc*A17 336
29644	cattle	1	4	2	2	22	9	17	620	18	16	40	16
29662	cattle	1	4	2	2	6	9	17	620	17	16	40	16.2
13261	beef offal	1	4	2	2	6	3	17	61	16.7	19.4	40	19
55706	cattle	1	4	2	16	6	3	17	60	17	16.4	40	18.1
55729	cattle	1	4	19	2	6	3	17	36	16.1	17.9	40	17
56054	cattle	2	4	2	2	6	4	17	81	17	17	40	17.2
Mean C_T values										16.2	17.1	40	17.2

8.3.6 SNP assay for clonal complex ST-206

(i) Design of primers and probes

The alleles *gln*A21 and *tkt*_1 had been identified to be predictive alleles for this clonal complex the SNPs used are summarised below.

Allele	SNPs
glnA21	C→T (18bp) A→G/T (33bp)
tkt_1	T→C (330bp)

The same allele *tkt*_1 and SNPs for the identification of ST-21 were used for identification of clonal complex ST-206; therefore the same uniplex assay could be used for both ST-21 and ST-206 (Section 8.3.2). A further uniplex assay was designed for the SNP on the *gln*A21 allele. For detection of the SNPs at 18bp and 33bp the allele sequence had to be extended by 50 bases to make it possible to fit in a primer before the location of the SNPs. This assay comprised the design of forward (21-43bp) and reverse (124-147bp) primers and an internal MGB probe labelled with VIC to detect the SNPs at 18bp and 33bp. These two SNPs were only 15bp apart therefore it was possible to use just one probe spanning both these SNPs (Table 8.1).

(ii) Results for reference isolates

The presence of the SNPs on the *gln*A21 allele was confirmed by a low C_T value of 19.2, which was present only within the ST-206 clonal complex assigned isolate (13263). This indicated a perfect match between the probe and target sequence therefore the presence of an C at 18bp and an A at 33bp. Throughout the remainder of the reference isolates results were negative with a high C_T value of 40. This indicated the presence of either one or two base changes at these positions therefore

no binding of the probe occurred and no detectable increase in fluorescence (Table 8.10).

The second SNP *tkt*_1 used for confirmation of this clonal complex was more widely distributed among the reference isolates (as described for ST-21 clonal complex) with three reference isolates showing the presence of this SNP including the ST-206 assigned isolate (13263) (Section 8.3.2). The presence of both these SNPs indicated the ST-206 clonal complex.

(iii) Application of the SNP assay for clonal complex ST-206 to larger sample numbers

By application of the assay to the larger sample set the assay was still valid and all the clonal complex ST-206 assigned isolates (n=5) were successfully identified in accordance with the results obtained by full MLST. Mean C_T values for each of these SNPs were 18.4 and 18.3. The use of the shared predictive allele *tkt*_1 with clonal complex ST-21 was also beneficial for this clonal complex. This facilitated the ultimate strategy, whereby fewer reactions could be used for the complete identification of all six clonal complexes.

It was established in the previous chapter that two predictive alleles were required for the accurate identification of a clonal complex. This was exemplified in the assays for ST-45 and ST-257 where one predictive allele had provided false identification of some clonal complexes. For these reasons two predictive alleles were used at the outset for the identification of this clonal complex. Clonal complex ST-206 was an additional clonal complex added into the strategy, which had not been described on

the Lightcycler. This enabled the complete identification of six of the major *C. jejuni* clonal complexes (Table 8.11).

Table 8.11 Results for all ST-206 assigned isolates (n=5) and consistent identification by SNP analyses

Strain	Source	Allelic Profile							ST	SNP result (C_T value)	
										*gln*A21 18, 33	*tkt*_1 336
48187	human stool	2	21	5	37	2	1	5	206	17.4	17
48766	chicken offal or meat	2	21	5	37	2	1	5	206	16	18.2
54558	human stool	2	21	5	37	2	1	5	206	21	19
55021	Human stool	1	21	5	37	2	1	5	232	20	18.7
13263	human stool	2	21	5	37	2	1	5	206	18	19
Mean C_T values										18.4	18.3

8.3.7 Assay for Clonal Complex ST-257

(i) Design of Primers and Probes for the Clonal Complex ST-257

Similarly the assay for ST-257 was modified from the Lightcycler assay to increase specificity. The allele *gly*A62 was used as in the Lightcycler assay but also with the additional allele *pgm*_4 and accompanying SNPs.

Allele	SNPs
glyA62	**C→T** (483bp)
pgm_4	**G→A/T** (162bp) **C→A/T/G** (165bp) **A→T/G** (168bp) **G→A** (171bp)

The assay comprised a duplex reaction to detect the one SNP on the *gly*A62 allele and the SNPs within the region 162-171bp on the *pgm*_4 allele. For the *gly*A62 assay forward (420-440bp) and reverse (492-517bp) and an internal probe labelled with VIC. Due to the positioning of the SNP on the *gly*A62 allele at the end of the allelic sequence the sequence had to be extended to allow for the reverse primer to fit. This was carried out in the same manner as described for the Lightcycler *gln*A4 assay design where the additional sequence was obtained from the sequence of the allele before it was trimmed to size. For the *pgm*_4 component of the duplex assay forward (163-181 bp) and reverse (269-295bp) primers were designed and an internal MGB probe labelled with FAM (Table 8.1)

(ii) Results for the Reference Isolates

Presence of the SNP on *gly*A62 at 483bp was confirmed by a C_T value of 18.3 for the ST-257 assigned isolate (13264) (Table 8.12). Throughout the remainder of the reference isolates the results were negative with a C_T value of 40 obtained in the

majority of cases indicating the presence of a base change to T, therefore no binding of the probe. The SNP used was not present in any other reference isolates.

Despite the good specificity of the SNP for the *glyA*62 allele, this allele was only 85% specific within the clonal complex and had been demonstrated to occur within other clonal complexes when the assay was applied to larger sample numbers (Section 7.3.5). Therefore the additional SNPs *pgm*_4 (162-171bp) were included to prevent the identification of false positive ST-257 clonal complex based upon *glyA*62 alone. The *pgm*_4 SNPs were widely distributed throughout the reference isolate set. As well as being present in the ST-257 clonal complex, isolate 13264, these SNPs were present in isolates 13261 (ST-61), 13260 (ST-52), 13265 (ST-354) 13266 (ST-362) and 13259 (ST-49). Despite the high distribution of these SNPs within other clonal complexes they did not occur within ST-45 clonal complex, as was the case for the *glyA*62 SNP alone, and therefore provided a suitable strategy in conjunction with *glyA*62 for ST-257.

(iii) Results for application of the assay for clonal complex ST-257 to larger sample numbers

The ST-257 duplex assay was applied to a larger sample set and similar low C_T values were obtained indicative of the ST-257 clonal complex. These results were confirmed with the assignment of ST-257 clonal complexes by the detection of SNPs in accordance with those obtained through full MLST. The data showed the necessity for the two alleles to confirm specificity for this clonal complex. For example the first predictive allele *glyA*62 was highly specific for the clonal complex, except for one isolate which assigned to clonal complex ST-45 (Sequence type 756)

(24,7,10,62,42,7,1), which also had an allele *gly*A62, in this case however, the *pgm*_4 SNP was not present, with the isolate having allele *pgm*_42, therefore not containing the SNPs within allele *pgm*_4. The duplex reaction provided a more convenient approach for identification of this clonal complex (Table 8.13).

Table 8.13 Results for all ST-257 assigned isolates (n=29) and consistent identification by SNP analyses

Strain	source	Allelic Profile							ST	SNP result (C_T value)	
										*glyA*62 483	*pgm_4* 162-171
47781	chicken offal or meat	2	2	4	62	4	5	6	367	16.7	14.7
47782	chicken offal or meat	2	2	4	62	4	5	6	367	16.9	14.7
48298	chicken offal or meat	9	2	4	62	4	5	6	257	18.4	14
48323	chicken offal or meat	9	2	5	62	4	5	1	717	16.5	14.9
48352	chicken offal or meat	9	2	4	62	4	5	6	257	16.7	15.5
48384	chicken offal or meat	2	2	4	62	4	5	6	367	16.4	14.6
48417	chicken offal or meat	2	2	4	62	4	5	6	367	17.7	17.4
48496	chicken offal or meat	4	2	4	62	4	5	6	366	16.8	14.2
48675	human stool	9	2	4	62	4	5	6	257	17.2	16.8
48742	chicken offal or meat	2	2	4	62	4	5	6	367	18.3	16
48782	chicken offal or meat	9	1	4	62	4	5	6	286	16.4	13.8
49407	human stool	9	2	4	62	4	5	6	257	17	14.7
49898	human stool	9	2	4	62	4	5	6	257	19.1	14.6
50820	chicken offal or meat	2	2	4	62	4	5	6	367	18.1	14.7
50830	chicken offal or meat	4	2	4	62	4	5	6	366	17	14.5
50995	human stool	9	2	4	62	4	5	6	257	19	13
50996	human stool	9	2	4	62	4	5	6	257	20	14.6
51538	human stool	9	2	4	62	4	5	6	257	18	14.7
51790	chicken offal or meat	4	2	4	62	4	5	6	366	17.2	14.5
51817	chicken offal or meat	2	2	4	62	4	5	6	367	17	16.4
52514	chicken offal or meat	1	2	4	62	4	5	6	737	16.5	14.1
52788	human stool	9	4	4	62	4	5	6	776	17.5	14.6
53896	human stool	8	2	5	62	4	5	1	584	17	14.3
54309	human stool	9	2	4	62	4	5	6	257	17.1	14.1
54471	human stool	9	10	4	62	4	5	6	777	16	14.3
54626	human stool	9	2	4	62	4	5	6	257	19.1	15.6
76665	human OB	9	2	4	62	4	5	6	257	18.2	15.4
12013	poultry	9	2	4	62	4	1	6	316	18.3	14.6
13264	human stool	9	2	4	62	4	5	6	257	19.4	16.9
Mean C_T values										17.5	14.9

8.3.8 Summary of SNP assay for all six clonal complexes

Table 8.14 represents a summary of results for all isolates tested, for all six clonal complexes and the comparison between the results obtained through the full MLST procedure and the SNP strategy developed in this chapter. Data was consistent for all the target clonal complexes ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257 and equal numbers of isolates were correctly identified by each method. This strategy based on the Taqman was an improvement over the strategy on the Lightcycler due to the additional alleles incorporated and the further development of the strategy to incorporate an additional clonal complex.

Several observations concerning the data in Table 8.1 were made regarding the distribution of positive results. For clonal complex ST-21 positive C_T values for each of the three specific SNPs were only seen together in the clonal complex ST-21 identified isolates, however positive results for each of these SNPs were seen individually within other clonal complex assigned isolates. Especially for the SNP *gln*A1 267, where positive C_T values were seen throughout many clonal complexes. Conversely for the *unc*A17 SNP a positive C_T value was only seen in the ST-61 assigned isolates.

It is encouraging that by comparing the Taqman results to those obtained on the Lightcycler (7.3.5) that the same distribution of positive results is seen, throughout the clonal complexes ST-21, ST-48 and ST-61. It is also clear to see that the results for the Taqman are more straightforward to interpret, the range of melting temperatures obtained by the lightcycler are more difficult to decipher. However the ranges of positive results across the non-target clonal complexes as in the case of the lightcycler

results suggest that it may be possible to determine other clonal complexes based upon positive results for the SNPs utilised here, which warrants further investigation.

8.3.9 Application of SNP assays for detection of specific *C. jejuni* strains

(i) Naturally contaminated meat samples

Further validation of the SNP strategy was required on other starting sample types. This was primarily to test the strategy for its use as a potential *C. jejuni* detection method and secondly to test the applicability of the technique for the detection of specific strains of *C. jejuni* by their clonal complex. The SNP strategy was tested on food samples, specifically meat samples that had been subjected to direct DNA extraction from the meat rinse (described in chapter 4).

In total 80% (n=10) of the culture confirmed meat samples were identifiable to their clonal complex when tested with the SNP assays (Table 8.15). The DNA samples from the meat samples were of highly variable DNA content and therefore produced a different range of C_T values compared to the assays used on DNA extracts from culture. Positive results as described earlier in this chapter when standardised quantities of DNA were used were in the range of 14-21. For the assays used on direct DNA extracts C_T values in the range of 30-38 were taken as being positive. High C_T values were assumed positive due to the quantitative feature of Taqman reactions, as described in chapter 4, higher C_T values are representative of lower volumes of DNA added into the reactions. Confirmation of clonal complex was carried out by full MLST directly from the cultured sample (as described in section 4), where identical clonal complex designations were determined in each case.

The remaining 20% (n=10) of culture confirmed meat samples tested were unable to be assigned a clonal complex by the SNP strategy. These were confirmed by full MLST to be assigned to other clonal complexes (ST-433 and ST-52), not

represented in the SNP scheme, therefore it was impossible for correct identification to be made.

In total, 57% (n=7) of the non-culture confirmed meat samples were identifiable with the SNP assays (Table 8.16). For three samples it was possible to confirm the clonal complex by full MLST, however one sample could not be determined. The chicken thigh sample had been determined by the SNP assays to belong to ST-206 clonal complex, however no successful MLST result could be obtained from this DNA extract despite numerous attempts. Reasons for this are difficult to explain. It could be the result of inhibitory components within the DNA extract from the chicken itself, which have interfered with the PCR, but this is unlikely when the SNP strategy was successful. Alternatively, the content of DNA in the sample could have influenced the result, whereby too low a volume of DNA was present in the sample to be effective in conventional PCR. The Taqman system incorporating MGB probes was more sensitive than conventional PCR therefore this may explain the absence of result by full MLST.

Current practice for detection of campylobacter within food samples involves traditional enrichment and culturing techniques, which can be time consuming. These are often not a true representation of the prevalence of the whole campylobacter population in the sample, due to selective media promoting the growth of some strains over others. Additionally, some types of campylobacters are more resistant to processing steps thus well-adapted strains are likely to be the ones most viable from cultured food samples. Limited universal methods exist for the direct detection of specific campylobacter strains from food samples (as described in

chapter 4). Consequently limited epidemiological studies have been carried out to determine the prevalence of different types of *C. jejuni* in food samples without using a prior enrichment or cultural step. Detection of specific types of campylobacters by PCR would be beneficial and potentially offer a different perspective on the distribution of *C.jeuni* strains throughout the food chain.

Table 8.15 Culture confirmed *C. jejuni* positive meat samples (ND not determined)

Source	Sample type	Clonal Complex from SNP strategy[1]	Confirmation by full MLST result[2]
Poultry-Chicken samples	Breast-1	ST-257	ST-257
	Legs-1	Negative	ST-52
	Legs (Halal)	ST-257	ST-257
	Diced breast (low fat)	ST-257	ST-257
	Legs (organic)	ST-21	ST-21
	Diced breast	ST-206	ST-206
	Goujons	ST-257	ST-257
	Chicken breast skin	Negative	ST-433
Poultry-Turkey samples	Steaks	ST-21	ST-21
Beef	Mince (low fat)	ST-61	ST-61

[1] Clonal complex determined through SNP assays directly on a MagNApure extract from meat sample.

[2] Clonal complex determined through full MLST following culture of meat sample.

Table 8.16 Non-culture confirmed *C. jejuni* positive meat samples (ND not determined)

Source	Sample type	MLST from SNP strategy[1]	Confirmation by full MLST result[2]
Poultry-Chicken samples	Breast (Free range)	ST-21	ST-21
	Thighs (boneless)	ST-257	ST-257
	Thighs (on the bone)	ST-206	ND
	Livers	Negative	ST-433
	Legs (frozen)	Negative	ST-42
	Fillets	ST-45	ST-45
Pork	Sausages –2	Negative	ST-403

[1] Clonal complex determined through SNP assays directly on a MagNApure extract from meat sample.

[2] Clonal complex determined through full MLST on same MagNApure extract.

(ii) Naturally contaminated water samples

DNA was isolated from water samples and confirmed *C. jejuni* positive (by Dr. S. Lai). These samples were also tested with the SNP assays to determine the clonal complex. In this instance no culture was available for confirmatory testing so full MLST was carried out on the DNA extracts.

In total 31.2% (n=16) of samples tested were identified by clonal complex by the SNP strategy. These included two samples identified as belonging to clonal complex ST-45 (E354 and E492), one belonging to clonal complex ST-257 (E491) and one potential mixture (E347), where the SNPs specific for both clonal complexes ST-45 and ST-206 were identified (Table 8.17). Of the remaining water samples, eight could not be determined with the SNP assay and were identified by full MLST to be clonal complexes ST-403 or ST-283. Additionally one sample (81) unidentified by the SNP strategy was unassigned to any clonal complex by full MLST.

Other observations within this data set revealed, as in the case of the meat samples, one sample (E357) that had been assigned clonal complex ST-45 by the SNP strategy, except no result was obtained by full MLST, despite numerous attempts. This was most likely to be a result of too small a quantity of DNA for the PCR to be effective. For two samples it was not possible to assign a clonal complex by either method. Despite these samples being confirmed as *C. jejuni* with the Taqman assay it was not possible to determine a clonal complex. These two strains could possibly be well-adapted strains of campylobacter displaying a great deal of sequence

divergence within the house keeping genes. For these reasons the SNPs would not be detectable and the primers used for full MLST would be ineffective.

Interestingly one sample was identified as being mixed ST-45 and ST-206 clonal complexes by the SNP analyses, but was confirmed as ST-45 by full MLST. Detection of mixtures could be a useful property of the SNP assays, which other techniques are unlikely to provide. With full MLST using an automated sequencer, mixtures are not easily detectable due to the automatic base calling, thus the higher peak on the chromatogram is usually recorded. If within a mixture, one clonal complex was more abundant than the other it would be expected that only the most abundant one would be detectable. Multiple campylobacter types would be likely in natural water samples due to contamination from farm animals thus detection of mixed infections warrants further investigation. It would be difficult to detect mixed infections by full MLST therefore the SNP strategy provides a useful additional insight and suggests that this strategy for the testing of environmental waters may be useful.

There are currently no studies confirming the distribution of clonal complexes of campylobacters within natural water samples, however data has suggested that the types found in cattle and poultry would be prevalent in natural water samples due to contamination of river water from farmland. This has possibly been confirmed here, where the clonal complexes ST-257, ST-45 and ST-206 associated with poultry (Colles *et al.*, 2003) and ST-403 from pigs (Manning *et al.*, 2003) have been identified. For these reasons the SNP strategy would be applicable for detection from naturally contaminated water samples. Nevertheless some of the MLST types

identified in the water samples by full MLST belong to some of the least common clonal complexes such as ST-283, which has only 28 assigned isolates on the database. This means that the feasibility for using the SNP strategy is limited, as it is not designed to detect the less common clonal complexes. The SNP strategy could be used as an initial screen to identify samples containing the most common clonal complexes leaving the less common types to be analysed by full MLST. The next step would be to apply the assays prospectively in a large study to test the validity of the results from water.

Detection of *C. jejuni* in water is important to determine routes of infection and for epidemiological studies. Conventional methods for detection of *C. jejuni* in water samples can be time consuming and lengthy, involving membrane filtration combined with enrichment and culturing techniques (Sails *et al.*, 2002). There is much controversy surrounding the use of selective media, particularly when some are reported to promote some strains over others, and consequently not give a true representation of campylobacter prevalence. Also the survival of different campylobacters in water is dependent on where they originated (Cools *et al.*, 2003) hence many strains do not survive well and are undetectable by culture due to being stressed or damaged and thus non-viable. This would suggest that PCR approaches are more applicable, especially if strains can be identified to a specific type. The results presented here imply that following large scale testing these assays would be applicable for screening of water samples for the most common clonal complexes, leaving a subset to be analysed by full MLST.

Table 8.17 Non-culture confirmed *C. jejuni* positive water samples (ND not determined)

Water sample	MLST from SNP strategy[1]	Confirmation by full MLST result[2]
E226	Negative	ST-403
E227	Negative	ST-403
65	Negative	ST-403
67	Negative	ST-403
81	Negative	Unassigned
82	Negative	ST-403
83	Negative	ST-403
E35	Negative	ST-403
E347 (1)	ST-45/ST-206	ST-45
E354 (2)	ST-45	ST-45
E357	ST-45	Negative
E490	Negative	ST-283
E491	ST-257	ST-257
E492	ST-45	ST-45
E493	Negative	Negative
E494	Negative	Negative

[1] **Clonal complex determined through SNP assays directly on a Prepman extract from water sample.**

[2] **Clonal complex determined through full MLST on same Prepman extract.**

8.4 General Discussion

Using the Taqman system in combination with carefully designed probes to target the informative SNPs a strategy has been designed and implemented which can identify the presence of SNPs and thus indicate if an unknown isolate assigns to one of the clonal complexes ST-21, ST-45, ST-48, ST-61, ST-206 or ST-257. The fluorogenic assays have been demonstrated to be reliable and accurate, since an obvious difference in the C_T value can be seen indicative of the presence of the SNP with results being obtained within a couple of hours. The probes used in all assays were effectively able to discriminate the presence of the SNP or the base change. This strategy as described for the Lightcycler allows for the rapid preliminary strain identification, which could inform epidemiological investigations. However the assays on the Taqman platform have potential advantages over the assays described for the Lightcycler system.

8.4.1 Advantages of the SNP assays on the Taqman platform

The strategy is similar to the one described in the previous chapter however the redesign for the Taqman allows for higher throughput, as 96 samples can be processed during one run. As is the case for the Lightcycler the Taqman platform also allows the possibility of duplexing reactions to create more convenient assays and reduces the time taken. The ability to multiplex more than two reactions together may be possible on the Taqman system in the near future, further Taqman MGB probe extensions are available such as the fluorescent label TET, however the Sequence Detection System Software is incapable of detecting three probes within one run. Additionally the Taqman probe and primer system is less complex with only

one probe required per SNP to be detected, compared to the Lightcycler set up where two probes are required which are more tedious to design and also more costly.

A large part of the time for the Lightcycler assay set up and implementation involved the optimisation of the reaction conditions in order to achieve robust melting peaks. This involved two main factors, (i) the optimisation of reaction components such as the $MgCl_2$, and (ii) the optimisation of reaction conditions especially the melting curve annealing temperature and time. Optimisation of $MgCl_2$ on the Lightcycler required multiple experiments to be carried out to determine the best working concentration within a range of 1-5mM, likewise the optimisation of the reaction conditions required multiple runs. Roche have recently introduced a new type of hybridisation probe mastermix that does not require optimisation of the $MgCl_2$, this would substantially reduce the time taken to optimise assays and reduce the amount of reagents used during this process.

Alternatively these assays on the Taqman have required very little optimisation in comparison to the Lightcycler. The Taqman MGB probes were designed to be optimal with the Taqman universal mastermix, and always worked first time when put onto the machine. This was an important consideration for extension of the strategy to other clonal complexes for the Taqman. The additional SNPs were added into the Taqman strategy to improve the specificity, which was easily achievable due to the small amount of optimisation required. In most cases the assays were adjusted with respect to annealing time and temperature to achieve the best discrimination for each particular SNP. However the addition of extra SNPs to the range of assays on

the Lightcyler would require more time consuming design of probes and assay optimisation.

Despite the Lightcycler machine having quick amplification of approximately one hour the set up and data analysis are more time consuming when compared to the Taqman. The Taqman benefits from using a 96 well format allowing for set up to be a relatively quick process by the use of multichannel pipettes or by the use of robotic liquid handling systems. The Lightcycler requires the mix of reagents and then transfer into glass capillaries, which require the application of plastic stoppers, which is a slow and fiddly process. Data analysis on the Lightcycler also is more time consuming compared to the data analysis on the Taqman.

Although the majority of the assays were carried out by using 25µl reaction volumes, some plates were set up by the use of a robotic liquid handling machine. For these plates reaction volumes were reduced by half to a total reaction volume of 12.5µl without loss of sensitivity, which made the Taqman assays more cost effective for high throughput screening. This was also described in a study by Morin *et al* (Morin *et al.*, 1999) where substantial reductions in assay volume were made without reducing the quality. It could be possible that further reductions in total assay volume could be made possibly to nanometer volumes to reduce costs, although this was not attempted during this investigation. Using this method the cost per test was £0.8 (for 25µl reaction volume) with the total amount for a screen for the six clonal complexes being £6.40. This is compared with a cost of £29.77 for full MLST per isolate in house. This makes the SNP MLST approach within acceptable financial limits for preliminary screening.

The ability of the Taqman assays to assign clonal complexes directly from extracted genomic DNA, rather than employing the first round PCR step that is required for the Lightcycler assay is another further potential benefit. This is likely to be due to the confirmed high specificity of the 5'→3' exonuclease activity of the *Taq* DNA polymerase within the Taqman chemistry even in the presence of a complex DNA background (Holland *et al.*, 1991). This means that the process is less time consuming due to the requirement for only one PCR step, less costly due to only one set of reagents being required, and less complicated to perform. This would make the strategy suitable for the implementation into a routine or reference laboratory where a scientist competent in real time PCR techniques could carry out the entire technique. Although the Lightcycler assay format only requires the one additional PCR step the process is more complicated and requires more complex data interpretation. For the purposes of routine surveillance then the Taqman assays would prove more cost and time effective, easier to use and enable more comparable results between laboratories.

8.4.2 Applicability for rapid detection

There are few techniques available for the real time detection and characterisation of *Campylobacter* species. As already mentioned, Lightcycler real time strategies have been described for example strategies for detection of different species (Logan, 2001; Stapleton *et al.*, 2003; Cheng & Griffiths, 2003) and Taqman assays have also been described (Nogva *et al.*, 2000a; Sails *et al.*, 2003a) but to date, no strategies applicable for PCR based intraspecies characterisation have been described other

than MLST, molecular fingerprinting techniques, PCR-ELISA techniques or sequencing strategies, none of which can be achieved in two hours.

The applicability of the SNP assays to other sample types including meat samples and environmental water samples is encouraging. This has been successfully achieved in conjunction with the MagNApure, which has provided a sensitive method for DNA extraction resulting in DNA of sufficient yields and quality to be successful in these assays. Obtaining naturally contaminated *C. jejuni* samples, which were viable, was problematical therefore only a limited sample set has been used. Nonetheless some interesting observations were made. Primarily it has been possible to detect specific strains of *C. jejuni* directly from contaminated meat and water samples. This strategy would be highly beneficial for rapid strain profiling and characterisation in outbreak investigations or for case cluster recognition to promote public health responses for early source identification. Currently there are no described methods for *Campylobacters*, which are rapid and enable the assigning of specific strain types direct from genomic DNA within 2 hours. Also observed is that it is possible to detect samples of mixed clonal complexes. Currently there are no PCR based methods that allow for the detection of multiple types. With most techniques one result is the only option especially with sequencing based strategies where mixed strains would be difficult to accurately detect.

8.4.3 Negative Aspects

The drawbacks in the SNP approach lie in that the data currently in the PubMLST database are only representative of the isolates that have been both typed by MLST and submitted, therefore as the database expands there is the potential for new alleles

to be missed. The process of assigning clonal complexes from SNP analysis may not suit all areas of research where MLST would potentially be considered. For population biology the full MLST data set would be required and alternative strategies for MLST such as the use of high-density DNA arrays as described for MLST of *Staphylococcus aureus* may be more applicable (van Leeuwen *et al.*, 2003). Additionally, assigning strains by clonal complex may not be sufficiently discriminatory in some cases and the sequence type may be required. A SNP MLST approach was described by Robertson *et al* (Robertson *et al.*, 2004). This was a bioinfomatics driven strategy where the informative SNPs used were representative of sequence types not clonal complexes, therefore the strategy was more discriminatory. However, the methods described were based on the Lightcycler and the detection of up to 14 SNPs was required for accurate identification of the sequence type, for these reasons the strategy in its current format would not be beneficial as a screening method for specific types. Also MLST, or the SNP approach as a technique may not be suitable in certain situations, a supplementary technique such as the sequencing of the short variable region of the *fla*A gene, was described as necessary for adequate discrimination in *C. jejuni* outbreak investigations (Sails *et al.*, 2003c).

The SNP strategy described here is limited by its detection of only six clonal complexes. This has been demonstrated in the use of the assays on water samples, where the samples belonging to the less common clonal complexes were obviously not detectable. For further work on water then the SNP strategy would need to be extended, or full MLST would still need to be applied to the remaining samples not detectable. In spite of this, the SNP strategy was never developed as a replacement typing technique to replace any existing methodologies; therefore its use as a first line

screen for the most common types is still highly beneficial for epidemiological studies, and this warrants further investigation. With the reduction in gastrointestinal infection by 20% being a major target for the Food Standards Agency and a key component to this being a reduction in the burden of infection due to *C. jejuni*. The SNP strategy described could be a suitable technique to address some of the issues relating to the lack of rapid and reproducible methods for campylobacter epidemiological studies.

Table 8.3 Example of results of Taqman SNP assays for ST-21 for the MLST Reference Strain Collection, identification of ST-21 assigned isolate by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP C_T Results			ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	ST-Complex	*glnA*1 108	*glnA*1 267	*tkt_*1 330	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	15.5	16	16	21
13257	Human stool	4	7	10	4	1	7	1	45	ST-45 complex	40	16	40	NA
13263	Human stool	2	21	5	37	2	1	5	206	ST-206 complex	40	16	18	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	40	40	40	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	40	40	20	NA
13264	Human stool	9	2	4	62	4	5	6	257	ST-257 complex	40	16	40	NA
13256	Human stool	1	2	3	4	5	9	3	42	ST-42 complex	40	16	40	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	40	17	40	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	40	17	40	NA
13265	Human stool	8	10	2	2	11	12	6	354	ST-354 complex	40	16	40	NA
13255	Human stool	1	3	6	4	3	3	3	22	ST-22 complex	40	40	40	NA
13266	Human stool	1	2	49	4	11	66	8	362	ST-362 complex	40	16	40	NA
13259	Human stool	3	1	5	17	11	11	6	49	ST-49 complex	16	16	40	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	40	40	40	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	40	17	40	NA

Table 8.5 Example of results of Taqman SNP assays for ST-45 for the MLST Reference Strain Collection, identification of ST-45 assigned isolate by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP C_T Results			ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	ST-Complex	*glt*A10 201	*glt*A10 225	*tkt_7* 138	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	40	15	40	NA
13257	Human stool	4	7	10	4	1	7	1	45	ST-45 complex	15.4	16	17.2	45
13263	Human stool	2	21	5	37	2	1	5	206	ST-206 complex	40	16	40	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	40	15	40	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	40	15.8	40	NA
13264	Human stool	9	2	4	62	4	5	6	257	ST-257 complex	40	15.5	40	NA
13256	Human stool	1	2	3	4	5	9	3	42	ST-42 complex	40	16	40	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	18	40	40	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	40	15.6	40	NA
13265	Human stool	8	10	2	2	11	12	6	354	ST-354 complex	40	16	40	NA
13255	Human stool	1	3	6	4	3	3	3	22	ST-22 complex	40	15	40	NA
13266	Human stool	1	2	49	4	11	66	8	362	ST-362 complex	40	15	40	NA
13259	Human stool	3	1	5	17	11	11	6	49	ST-49 complex	40	15	40	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	40	15	18	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	40	15	40	NA

Table 8.7 Example of results of Taqman SNP assays for ST-48&61 for the MLST Reference Strain Collection, identification of ST-48 & 61 assigned isolates by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP C_T Results				ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	ST-Complex	*glnA*4 18	*glnA*4 202	*unc*A5 186,189	*unc*A17 336	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	40	40	14.3	40	NA
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	40	40	40	40	NA
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	15.9	40	15.17	40	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	16.7	19.4	40	16	61
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	16	19	14.3	40	48
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	40	40	40	40	NA
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	40	40	40	40	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	40	40	40	40	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	40	40	40	40	NA
13265	human stool	8	10	2	2	11	12	6	354	ST-354 complex	40	17	40	40	NA
13255	human stool	1	3	6	4	3	3	3	22	ST-22 complex	40	17	40	40	NA
13266	human stool	1	2	49	4	11	66	8	362	ST-362 complex	40	40	40	40	NA
13259	human stool	3	1	5	17	11	11	6	49	ST-49 complex	40	40	40	40	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	20	40	15	40	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	40	40	40	40	NA

Table 8.10 Example of results of Taqman SNP assays for ST-206 for the MLST Reference Strain Collection, identification of ST-206 assigned isolates by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP C_T Results		ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	ST-Complex	*glnA*21 18,33	*tkt_*1 330	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	40	16	NA
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	40	40	NA
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	19.2	18	206
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	40	40	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	40	20	NA
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	40	40	NA
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	40	40	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	40	40	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	40	40	NA
13265	human stool	8	10	2	2	11	12	6	354	ST-354 complex	40	40	NA
13255	human stool	1	3	6	4	3	3	3	22	ST-22 complex	40	40	NA
13266	human stool	1	2	49	4	11	66	8	362	ST-362 complex	40	40	NA
13259	human stool	3	1	5	17	11	11	6	49	ST-49 complex	40	40	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	40	40	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	40	40	NA

Table 8.12 Example of results of Taqman SNP assays for ST-257 for the MLST Reference Strain Collection, identification of ST-257 assigned isolate by MLST and SNP analyses.

Isolate	Source	Allelic Profile									SNP C_T Results		ST-complex Assigned by SNP Analysis
		aspA	*glnA*	*gltA*	*glyA*	*pgm_*	*tkt_*	*uncA*	ST	ST-Complex	*glyA*62 483	*pgm_*4 165	
13254	Beef offal	2	1	1	3	2	1	5	21	ST-21 complex	40	40	NA
13257	human stool	4	7	10	4	1	7	1	45	ST-45 complex	40	40	NA
13263	human stool	2	21	5	37	2	1	5	206	ST-206 complex	40	40	NA
13261	Beef offal	1	4	2	2	6	3	17	61	ST-61 complex	40	16	NA
13258	Lamb offal	2	4	1	2	7	1	5	48	ST-48 complex	40	40	NA
13264	human stool	9	2	4	62	4	5	6	257	ST-257 complex	18.3	18.4	257
13256	human stool	1	2	3	4	5	9	3	42	ST-42 complex	40	40	NA
13262	Sand	17	2	8	5	8	2	4	177	ST-177 complex	40	40	NA
13260	Lamb offal	9	25	2	10	22	3	6	52	ST-52 complex	40	14.7	NA
13265	Human stool	8	10	2	2	11	12	6	354	ST-354 complex	40	14.6	NA
13255	Human stool	1	3	6	4	3	3	3	22	ST-22 complex	40	40	NA
13266	Human stool	1	2	49	4	11	66	8	362	ST-362 complex	40	14.8	NA
13259	Human stool	3	1	5	17	11	11	6	49	ST-49 complex	40	14.7	NA
13268	Unknown	10	27	16	19	10	7	5	403	ST-403 complex	40	40	NA
13267	Unknown	7	17	5	2	10	6	3	353	ST-353 complex	40	40	NA

Table 8.14 Summary of results for 221 isolates showing the mean C_T values for each assay and those indicative of specific clonal complexes. ND-not determined by this assay. [1] Clonal complex determined by full MLST (Dingle *et al.*, 2001). [2]Clonal complex determined by detection of SNPs on the Taqman. *Any combination of C_T values possible. (Shaded areas indicating positive C_T values, indicative of the clonal complex).

C. jejuni full MLST clonal complex[1] (number of isolates)	Mean C_T value obtained for each SNP														Designated Clonal Complex by SNPs[2] (number of isolates)
	ST-21			ST-45			ST-257		ST-61 & 48				ST-206		
	glnA1 108	*glnA1* 267	*tkt_1* 330	*gltA10* 201	*gltA10* 225	*tkt_7* 138, 141	*glyA62* 483	*pgm_4* 162-171	*glnA4* 18	*glnA4* 202	*uncA5* 186, 189	*uncA* 17 336	*glnA* 21 18, 33	*tkt_1* (as for ST-21)	
ST-21 (51)	16	16.4	19	16/40	17/40	40	40	40	40	40	16/40	40	40	19	ST-21 (51)
ST-257 (29)	40	15/40	40	40	16	40	17.5	14.8	16/40	17/40	40	40	40	40	ST-257 (29)
ST-61 (3)	40	40	40	40	15	40	40	16/40	17.2	19.1	40	19.7	40	40	ST-61(3)
ST-48 (21)	40	15/40	14	40	16	16/40	40	40	16.7	18.2	17.6	40	40	14	ST-48 (21)
ST-45 (28)	40	16/40	40	17.1	16	16	17/40	40	16/40	17/40	40	40	40	40	ST-45 (28)
ST-206 (5)	40	16/40	15	40	16	40	40	40	16/40	17/40	40	40	18	15	ST-206 (5)
ST-49 (4)	16	15	40	16/40	17/40	40	40	16	40	40	16/40	40	40	40	ND
ST-22 (7)	40	40	40	16/40	15/40	40	40	40	40	16/40	40	40	40	40	ND
ST-177 (2)	40	16	40	16	40	40	40	40	40	40	40	40	40	40	ND
ST-353 (4)	40	16	40	40	17	40	40	16/40	40	40	40	40	40	40	ND
ST-42 (10)	40	16	40	40	16.1	40	40	40	40	40	40	40	40	40	ND
ST-52 (7)	40	16	40	40	17	40	40	16/40	40	40	18/40	40	40	40	ND
ST-354 (10)	40	15	40	40	16/40	16/40	40	16	16/40	40	40	40	40	40	ND
ST-403 (4)	40	17/40	40	16	40	17/40	40	17/40	16/40	40	16/40	40	16/40	40	ND
ST-362 (1)	40	16	40	40	17	40	40	16/40	40	40	40	40	40	40	ND
Other and unassigned(35)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	ND

Chapter 9

Final discussion and future work.

Chapter 9

Final discussion and future work

9.1 Final discussion

At the start of this investigation molecular fingerprinting methods were the current methods for molecular characterisation of campylobacter strains, particularly in outbreak investigations. This thesis initially describes preliminary investigations determining the suitability of the molecular fingerprinting techniques (SAFLP and PFGE) for use in detection and characterisation of campylobacters. Both PFGE and SAFLP were determined to be successful methods for the recognition of epidemiological groupings of campylobacter as demonstrated here in the context of an outbreak investigation. Additionally, equal discrimination was achieved with both techniques suggesting that SAFLP, due to being the less complex and time consuming technique could be used in circumstances where PFGE would usually be considered. In spite of this, PFGE is still considered the "Gold standard" genotyping technique and with many established schemes, such as Pulsenet and Campynet it is unlikely that SAFLP will become as widely used for campylobacter epidemiological studies. Moreover with newer, PCR based methodologies and "next generation" sequence based typing methods in existence, then molecular fingerprinting methods are likely to become obsolete within the next few years.

Investigations to determine the suitability of SAFLP as a detection method were carried out. Adaptations to the method were made to improve turnaround time, however the technique was still restricted by the starting material. DNA within a specific range (0.1-1μg/μl) was required and the addition of DNA outside of this resulted in failure. This

verified the unsuitability of the SAFLP technique to be used in detection studies where unknown and potentially smaller volumes of DNA would be present. As described in later sections of this thesis, real time PCR techniques such as the Taqman assay for speciation were capable of detection of quantities of DNA as small as 20fg per PCR reaction, therefore the AFLP technique could not reach the required sensitivity when compared to newer technologies. FTA filters were tested with the AFLP technique but loss of DNA integrity occurred, resulting in the creation of different AFLP patterns from the same starting material. These inconsistent results produced meant that the FTA filters were not investigated further.

With the availability of real time PCR platforms, the next stage in the investigation was to devise a real time PCR assay for the detection of *C. jejuni* and *C.coli.* Current methods for speciation of the two commonest campylobacters were based upon phenotypic tests or conventional PCR methodologies with variable validity. The Taqman system offered potential advantages with increased throughput, sensitivity and accuracy. An assay was designed which was able to specifically detect the two species based upon two genes, the *map*A gene which was specific for *C. jejuni* and a region of the *ceu*E gene which was specific for *C.coli.* The assay was successfully developed and was optimised to reduce costs and improve sensitivity. The assay was confirmed as specific for the two species and highly sensitive, where 12.5 genome copies per PCR reaction were detectable for the *map*A gene and 174 genome copies detectable for the *ceu*E gene. The validity of assay was tested on a larger scale, as a front line molecular speciation test within the Campylobacter Reference Laboratory where successful speciation was possible for 97.6% (n=6015) of isolates. A number of interesting observations were made from this sample set. Primarily, mixtures of *C. jejuni* and *C.coli*

were detectable; these were further investigated where some were genuine mixed cultures, yet some were determined to be potential hybrid *C. jejuni/C.coli* strains where separation was unachievable. Further confirmation of mixtures was obtained through sequencing of both genes from each isolate. With no consistent results achieved through this and by other characterisation methods, it was concluded that these were representative of a subset of strains, which may have adapted possibly due to different environmental conditions. Coinfection can occur with *Campylobacter* species (Richardson *et al.*, 2002) and there is a growing amount of evidence that mixed infections of either species are significant features of campylobacter infection (Englen & Fedorka-Cray 2002).

Several other observations regarding the results from the large-scale study were made. In total 2.1% of strains were nontypeable by the Taqman assay, however these were typeable by other methods. Few studies have been described where one method for speciation had been applied to as many isolates (n=6015). It is possible that the application of any PCR technique to this number of strains would result in a similar number of strains failing. With a highly variable organism such as campylobacter, no typing technique is expected to be 100% successful. Single Nucleotide Polymorphisms will occur in DNA as a result of insertions and deletions due to environmental stress and adaptation. Interestingly, many of the isolates, which failed, were of an environmental source. Since the *ceuE* gene is postulated as being a potential virulence gene due to its involvement in iron sequestration, it is possible that divergence has occurred within the gene, away from the published sequence as a result of environmental influences.

A number of DNA extraction methods were evaluated with this assay to assess sensitivity, compare cost and set up time. The two methods on the MagNApure were tested against the method of DNA preparation by boiling (cell lysates) using bacterial cultured cells. Optimal sensitivity was obtained with the MagNApure Bacterial DNA extraction Kit where the minimal levels of *C. jejuni* detectable were $1x10^1$ CFU/ml. The cell lysate method resulted in the highest minimum recovery rate of $1x10^4$ CFU/ml, however, the low cost and short preparation time was advantageous for rapid throughput and this was used in future investigations from culture.

Direct detection was also demonstrated using the real time PCR methods. Very limited studies have been described for direct detection of *Campylobacter* species from starting sample types such as meat rinses. There are many problems surrounding the use of PCR detection from food samples, owing to the interfering components within the food sample matrix, which can adversely effect the PCR reaction. Three methods were tested for direct DNA extraction from artificially contaminated meat rinse and milk samples. These included the MagNApure, Prepman and the Bugs 'n Beads systems, which were investigated by sensitivity, cost and time taken. The MagNApure system was determined to be the most sensitive system for all the sample types tested where detection limits of $1x10^1$ to $1x10^3$ CFU/ml were obtained across the whole range of sample types tested. The other two methods tested did not demonstrate optimum reproducibility with all the sample types tested and there were a few negative aspects to the systems. The Prepman system did not work optimally with some of the food samples due to the excess fat which accumulated on the supernatant, however detection limits obtained were similar to those of the Bugs'n Beads system. Furthermore, the reproducibly of the Bugs 'n Beads system was not optimal, some beads were lost during

the pipetting stages, and the interference of the beads with the fluorophores of the Taqman chemistry was a consideration. The MagNApure resulted in the highest recovery rates from all the sample types tested, in conjunction with fast extraction times and was therefore used in all subsequent investigations.

Direct detection was also demonstrated on naturally contaminated samples. Due to the difficulties in obtaining naturally contaminated samples, only meat rinse samples were investigated. It was possible to confirm the presence of *C. jejuni* or *C.coli* in 48% (n=35) of samples by the Taqman assay, of which 28% were culture confirmed. Direct detection with this assay would aid epidemiological studies where a result could be obtained within three hours of receiving a food sample. Present methods require that the potentially contaminated food sample be cultured, and then identified either by phenotypic methods or conventional PCR. Many positive samples by Taqman were not culture confirmed, suggesting that the cells were damaged or stressed. This is in accordance with other studies where it has been reported that traditional culturing results in significant under reporting of potentially infectious *C. jejuni* or *C.coli.* (Atabay & Corry, 1997; Nogva *et al.*, 2000a). The concept of undetectable campylobacters by culture has been demonstrated in transmission studies where no detectable growth by culture has been found, but damaged or stressed cells have been described (Pearson *et al.*, 1993). Large direct detection studies utilising an assay such as this would provide a different insight into the prevalence of the two species of campylobacter within food samples.

This assay has provided a complementary tool for the rapid speciation of *C. jejuni* and *C.coli*, which can also be applied to the Lightcycler real time system and used as a

detection method from potential sources of infection including meat samples. This assay is still being used to date for the front line speciation of isolates received into the reference laboratory and used by other campylobacter research groups within other laboratories in the UK.

The second section of the thesis describes the strategy developed for the more comprehensive molecular detection and intraspecies characterisation of specific *C. jejuni* types. The process involved in the design of assays for the rapid characterisation of *C. jejuni* isolates, based upon their MLST clonal complex is described for the identification of predictive alleles for six target clonal complexes ST-21, ST-45, ST-48, ST-61, ST-206 and ST-257. The identification of the informative SNPs within these alleles and the evolution of this data into rapid and accurate assays is outlined. The allelic discrimination assay formats investigated included real time PCR using both the Taqman and Lightcycler platforms demonstrating maximum instrument and assay flexibility.

The first stage of this process was the identification of predictive alleles from the data currently within the *C. jejuni* MLST database. Detailed assessments of all the data were made to determine predictive alleles within six target clonal complexes, which could be used for identification purposes. It was possible to determine at least two predictive alleles for each clonal complex, and informative SNPs, which enabled the subsequent characterisation. The strategy based upon the Lightcycler was a preliminary test to determine if the strategy would be feasible. Assays were developed for five target clonal complexes, where two predictive alleles were used for the identification of clonal complexes ST-21, ST-48 and ST-61 and one predictive allele used for clonal complexes

ST-45 and ST-257. The assays were tested with the MLST reference collection (Wareing *et al.*, 2003) where results were consistent with the full MLST designation. Additionally they were tested with a larger panel of strains encompassing strains from different sources. For clonal complexes ST-21, ST-48 and ST-61 all the isolates were correctly assigned. However, results for the other two clonal complexes ST-45 and ST-257 were not 100% specific due to these assays being based upon only one predictive allele. For these reasons it was determined that two predictive alleles were required to achieve the adequate specificity required for these clonal complexes.

Several other observations were made regarding development of these assays on the Lightcycler. The inability to use genomic DNA in the assay was of concern, and the generation of first round PCR products increased the overall assay time by an additional three hours. Also the low throughput (32 samples) of the Lightcycler machine and the time consuming data analysis made the process slow despite its fast cycling times. Another consideration for the strategy on the Lightcycler was the extensive optimisation required to determine optimal reagent conditions and cycling parameters. For these reasons the SNP strategy was not developed further on the Lightcycler and the Taqman real time PCR platform was considered for further development and extension of the strategy.

The SNP assays on the Taqman were developed using a similar strategy to those based upon the Lightcyler, however the strategy was expanded to include additional predictive alleles for the clonal complexes ST-48 and ST-257, which were not sufficiently specific on the Lightcycler. Additionally the strategy was extended to incorporate the clonal complex ST-206. The Taqman primer and probe format was slightly different to that

described on the Lightcycler, despite the identical SNPs being used for detection purposes. The Lightcycler assays required that the primers were located in relatively conserved areas to allow for amplification to occur across all the locus alleles. However, with the Taqman system, the primers could be located within divergent regions and consequently the assays on the Taqman were more specific for the particular allele. The resultant strategy was able to assign effectively six target clonal complexes using a total of ten predictive alleles encompassing 21 SNPs. Costs were kept to a minimum by the use of shared predictive alleles for the clonal complexes ST-21 and ST-206 (*tkt*_1) and ST-61 and ST-48 (*gln*A4).

Specificity of the assays for *C. jejuni* was confirmed by testing with NCTC strains of other *Campylobacter* species, related organisms and other enteric pathogens, also the specificity of the assays for the six target clonal complexes was confirmed by use of the MLST reference isolates (Wareing *et al.*, 2003). With adequate specificity obtained, the SNP strategy was applied to a larger strain set, where all isolates belonging to the six target clonal complexes were successfully characterised. The sample set used encompassed strains chosen randomly, with equal numbers obtained from retail poultry, and cases of human enteritis. The distribution of clonal complexes seen was consistent with those described elsewhere (Dingle *et al.*, 2001), with the majority of isolates assigned to clonal complexes ST-21, and smaller numbers of strains assigning to other clonal complexes.

The SNP strategy was also tested for its feasibility as a direct detection method from potential sources of infection. The meat rinse samples described in chapter 4, which were confirmed *C. jejuni* positive, were subjected to the SNP assays for further

characterisation. The direct DNA extracts were used and successful confirmation of isolates belonging to the six target MLST clonal complexes was achieved. Where possible, confirmatory testing by full MLST was carried out on the culture, or alternatively full MLST was applied to the direct DNA extract to confirm the SNP assays were correct.

Also available for testing were *C. jejuni* positive, naturally contaminated water samples where campylobacter genomic DNA had been extracted by use of the Prepman kit. The SNP strategy was applied to these samples where 32.1% of the samples were successfully assigned a clonal complex, which was confirmed by full MLST. All the remaining water samples belonged to other clonal complexes, which were not included in the SNP scheme. These results demonstrated the usefulness of the SNP strategy as a preliminary screen. It was possible to identify the isolates belonging to the six target clonal complexes, which were confirmed through full MLST, leaving fewer samples to be characterised by the full technique.

The techniques described in this thesis enable the accurate detection of *C. jejuni* and *C.coli* and further *C. jejuni* strain characterisation by MLST clonal complex. The techniques developed have been demonstrated to be functional over a range of potential sources of campylobacter infection and compatible with different starting materials. For example, the Taqman speciation assay and SNP strategy can both be applied to genomic DNA, cell lysates and direct DNA extractions from sample types such as charcoal transport swabs, artificially contaminated meat and milk samples and naturally contaminated meat and water samples. All these assays can be performed in less than

three hours directly from DNA (or cell lysate) and in less than five hours directly from a charcoal swab/meat sample/water sample in conjunction with the MagNApure.

9.2 Suggestions for Future Work

The true incidence of the less common campylobacters in food and animals may be underestimated, as most isolation methods have been developed to isolate the commonest thermophillic campylobacters, *C. jejuni* or *C. coli* (Corry *et al.*, 1995). It would be useful to extend the Taqman speciation assay described here to incorporate genes specific to other species of *Campylobacter* and develop a multiplex assay, which could be used to identify more than just the two main *Campylobacter* species. This could involve the use of genes specific to other species of *Campylobacter* such as the *sap*A locus of *C. fetus*, which has no known homolgue in other *Campylobacter* species. Alternatively divergent regions within the same gene could be utilised, such as the *ceu*E gene described in this project or regions of the 16S rRNA gene. To date, no real time Taqman assays have been described for the other less common species of *Campylobacter* due to the low current isolation rates for these organisms. However, the increased specificity and accuracy of real time PCR assays allow for more valid testing to be carried out and could also overcome some of the problems of inhibition by selective medias. It would additionally be useful to incorporate a target for *Arcobacter* into this type of real time PCR strategy as *Arcobacter* species have been described to be abundant on chicken carcasses (Atabay & Corry, 1997).

A number of observations were made throughout the development of the SNP assays, which could be investigated further. The SNP strategy could be extended to incorporate other clonal complexes. It would not be feasible to extend the strategy to include all

clonal complexes, however the inclusion of assays for clonal complex ST-22 that has been associated with Guillain Barré and Miller Fisher syndromes (Dingle *et al.*, 2001) would benefit epidemiological studies. The data presented in this project suggests that extension of the strategy to other infrequent clonal complexes could be feasible by the presence of possibly one predictive SNP. The strategy could be of a hierachical nature whereby a preliminary screen could be carried out for the commonest clonal complexes as described here, followed by a second screen for some of the less common clonal complexes. There is also the potential for the SNP scheme to be extended to other species especially with the *C. coli* MLST scheme under development. It would be useful to add specific SNPs for some *C.coli* specific clonal complexes into the strategy to provide greater coverage.

From a public health perspective the application of the SNP assays to clinical samples, especially faecal samples would be beneficial. This could potentially be carried out in conjunction with a specific DNA extraction kit to allow for direct detection from the faecal material without the requirement for a culturing step. This could provide an interesting insight into the presence of different campylobacters, without the prohibitive effects of culture conditions. There are currently no other typing strategies which can be used in this manner and which enable the detection of specific *Campylobacter*s directly from faecal material.

There are many aspects relating to the distribution of clonal complexes and source and the relationships within the clonal complexes. With the large ubiquitous distribution of clonal complex ST-21 assigned isolates and the suggestion that they are well adapted for long term survival in a number of hosts (Colles *et al.*, 2003), it would be interesting

to use the SNP approach to identify ST-21 isolates and carry out viability testing. Strain relationships within the ST-21 clonal complex could also be investigated by other methods. Whole genome microarrays have been described for *C.jejuni* (Dorrell *et al.*, 2001), which could be used to investigate the presence or absence of other genes within the isolates assigned to this clonal complex.

The SNP assays were developed on two different real time PCR platforms, where only two reactions could be multiplexed. With the advances in real time PCR chemistries it would now be possible to multiplex more reactions together and therefore enable a greater number of SNPs to be detected within one PCR reaction. Further development on alternative platforms may offer potential advantages. Of particular value would be implementation of the SNP assays on the Pyrosequencer. This would enable quicker detection of SNPs (30 minutes for 96 samples) and the capability of multiplexing more assays together. As the SNP strategy is based on the detection of specific SNPs which act as markers, it could also be feasible to adapt the strategy to a microarray format, offering the potential for the addition of more SNPs and other potential targets including markers for species or antibiotic resistance. With the development of a MLST scheme for *C.coli*, and concomitant requirement to detect a greater number of SNPs, then it may be more pragmatic to adopt this format.

A further interesting observation is the ability to detect mixed infections of multiple *C. jejuni* types. Existing methods for characterisation of *C. jejuni* types only offer the capacity for the recogntion of one specific type. This is especially the case for sequencing reactions where the more predominate type would mask the presence of a less predominant one within the same sample. The SNP assays demonstrated here allow

the recognition of more than one type of campylobacter, due to use of multiple targets specific for each clonal complex and the increased sensitivity of the Taqman system, where a less predominant type would also be detected. This approach would be especially useful in large scale screening studies directly from clinical samples such as faeces where mixed infections could be possible.

The potential rapidity of the SNP assays, low cost and the robustness of the chemistry as demonstrated in this investigation surpass other methods for characterisation of *C. jejuni*. For these reasons the assays could be applied to potential sources of campylobacters where there is currently relatively little information. For example, in biofilms, naturally contaminated water samples and bottled/drinking water samples, where there is much speculation regarding the presence or absence of *C. jejuni*. The contribution of different types of *Campylobacter* in different environments would provide possible insights into niche adaptation and prevalence and ultimately inform epidemiological studies.

References.

References

Achtman, M. & Suerbaum, S. 2001, "Population Structure of *Helicobacter* pylori and other Pathogenic Bacterial Species," in *Helicobacter pylori: Molecular and Cellular Biology*. Horizon Scientific Press, Wymondam, UK, pp. 311-318.

Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A., and Carniel, E. 1999, *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*, Proc. Natl. Acad. Sci.U.S.A, vol. 96, no. 24, 14043-14048.

Ajayi, T., Allmond, L. R., Sawa, T., and Wiener-Kronish, J. P. 2003, Single-nucleotide-polymorphism mapping of the *Pseudomonas aeruginosa* type III secretion toxins for development of a diagnostic multiplex PCR system, J. Clin. Microbiol., vol. 41, no. 8, 3526-3531.

Al Rashid, S. T., Dakuna, I., Louie, H., Ng, D., Vandamme, P, Johnson, W., and Chan, V. L. 2000, Identification of *Campylobacter jejuni, C.coli, C.lari, C.upsaliensis, Arcobacter butzleri* and *A.butzleri-like* species based on the *glyA* gene, J. Clin. Microbiol., vol. 38, no. 4, 1488-1494.

Allerberger, F., Al Jazrawi, N., Kreidl, P., Dierich, M. P., Feierl, G., Hein, I., and Wagner, M. 2003, Barbecued chicken causing a multi-state outbreak of *Campylobacter jejuni* enteritis, Infection, vol. 31, no. 1, 19-23.

Allmann, M., Hofelein, C., Koppel, E., Luthy, J., Meyer, R., Niederhauser, C., Wegmuller, B., and Candrian, U. 1995, Polymerase chain reaction (PCR) for detection of pathogenic microorganisms in bacteriological monitoring of dairy products, Res. Microbiol., vol. 146, no. 1, 85-97.

Alm, R. A., Guerry, P., and Trust, T. J. 1993, Distribution and polymorphism of the flagellin genes from isolates of *Campylobacter coli* and *Campylobacter jejuni*, J. Bacteriol., vol. 175, no. 10, 3051-3057.

Altekruse, S. F. and Tollefson, L. K. 2003, Human campylobacteriosis: a challenge for the veterinary profession, J. Am. Vet. Med. Assoc., vol. 223, no. 4, 445-452.

Ang, C.W., Noordzij, P.G., de Klerk, M.A., Endtz, H.P., van Doorn, P.A., and Laman, J.D. 2002, Ganglioside mimicry of Campylobacter jejuni lipopolysaccharides determines antiganglioside specificity in rabbits, Infect. Immun. vol. 70, no. 9, 5081-5085.

Anon. 1997 Relative Quantitation of Gene Expression. ABI PRISM 7700 Sequence Detection System. User Bulletin #2, pp.3-10

Anon. 1998 PHLS Methods For Water Detection of Campylobacter species by membrane filtration. Standard Method : W8. Technical Services PHLS HQ. London UK.

Anon 2000, A Report of the Study of Infectious Intestinal Disease in England, The Food Standards Agency, Crown Copyright. Her Majesty's Stationery Office, Norwich. UK, pp. 120-125

Anon 2001, Campynet Website, www.svs.dk/campynet. Accessed 08/04/01

Anon 2003 Principles and Applications: Basic Applications for the Lightcycler System. http://www.lightcycler-online.com . Accessed 22/11/03

Anon 2004 "4th Symposium on Molecular Serotyping" 4/03/04 Centres for Disease Control and Prevention. Atlanta. USA.

Applied Biosystems 2001 SDS 7700 Training Materials, Section 1, pp 1-11.

Applied Biosystems 2002, Taqman MGB Probes, SDS News, vol. 17, no. B, 1-2.

Aspinall, G. O., McDonald, A. G., Raju, T. S., Pang, H., Mills, S. D., Kurjanczyk, L. A., and Penner, J. L. 1992, Serological diversity and chemical structures of *Campylobacter jejuni* low-molecular-weight lipopolysaccharides, J. Bacteriol., vol. 174, no. 4, 1324-1332.

Atabay, H. I. and Corry, J. E. 1997, The prevalence of campylobacters and arcobacters in broiler chickens, J. Appl. Microbiol., vol. 83, no. 5, 619-626.

Atabay, H. I., Corry, J. E., and On, S. L. 1997, Isolation and characterization of a novel catalase-negative, urease-positive *Campylobacter* from cattle faeces, Lett. Appl. Microbiol., vol. 24, no. 1, 59-64.

Attenborough, M. and Matthews, K. R. 2000, Food safety through the meat supply chain, Symp. Ser. Soc. Appl. Microbiol. no. 29, 144S-148S.

Baig, B. H., Wachsmuth, I. K., and Morris, G. K. 1986, Utilization of exogenous siderophores by *Campylobacter* species, J. Clin. Microbiol., vol. 23, 431-433.

Bang, D. D., Wedderkopp, A., Pedersen, K, and Madsen, M. 2002, Rapid PCR using nested primers of the 16S rRNA and the hippuricase (*hip*O) genes to detect *Campylobacter jejuni* and *Campylobacter coli* in environmental samples, Mol. Cell. Probes, vol. 16, 359-369.

Benoit, P. W. and Donahue, D. W. 2003, Methods for rapid separation and concentration of bacteria in food that bypass time-consuming cultural enrichment, J. Food. Prot., vol. 66, no. 10, 1935-1948.

Bereswill, S. and Kist, M. 2003, Recent developments in Campylobacter pathogenesis, Curr. Opin. Infect. Dis., vol. 16, no. 5, 487-491.

Best, E. L., Powell, E. J., Swift, C., Grant, K. A., and Frost, J. A. 2003, Applicability of a rapid duplex real-time PCR assay for speciation of *Campylobacter jejuni* and *Campylobacter coli* directly from culture plates, FEMS Microbiol. Lett., vol. 229, no. 2, 237-241.

Black, R. E., Levine, M. M., Clements, M. L., Hughes, T. P., and Blaser, M. J. 1988, Experimental *Campylobacter* infection in humans, J. Infect. Dis., vol. 157, 472-479.

Blaser, M. J. 1995, "Campylobacter and Related Species," in *Principles and Practice of Infectious Disease*, G. L. Mandell, J. E. Bennett, & R. Dolin, eds., Churchill Livingstone Inc, New York, pp. 1948-1956.

Blaser, M. J., Hardesty, H. L., Powers, B., and Wang, W. L. 1980, Survival of *Campylobacter fetus* subsp *jejuni* in biological milieus, J. Clin. Microbiol., vol. 11, 309-313.

Boddicker, J. D., Ledeboer, N. A., Jagnow, J., Jones, B. D., and Clegg, S. 2002, Differential binding to and biofilm formation on, HEp-2 cells by *Salmonella enterica* serovar *Typhimurium* is dependent upon allelic variation in the *fim*H gene of the *fim* gene cluster, Mol. Microbiol., vol. 45, no. 5, 1255-1265.

Bolton, F. E., Holt, A. V., and Hutchinson, D. N. 1984a, Campylobacter biotyping scheme of epidemiological value, J. Clin. Path., vol. 37, 677-681.

Bolton, F. J., Coates, D., Hutchinson, D. N., and Godfree, A. F. 1987, A study of thermophilic campylobacters in a river system, J. Appl. Bacteriol., vol. 62, no. 2, 167-176.

Bolton, F. J., Hutchinson, D. N., and Coates, D. 1984b, Blood-free selective medium for isolation of *Campylobacter jejuni* from feces, J. Clin. Microbiol., vol. 19, no. 2, 169-171.

Bolton, F. J., Hutchinson, D. N., and Coates, D. 1986, Comparison of three selective agars for isolation of campylobacters, Eur. J. Clin. Microbiol., vol. 5, no. 4, 466-468.

Bolton, F. J. and Robertson, L. 1982, A selective medium for isolating *Campylobacter jejuni/coli*, J. Clin. Pathol., vol. 35, no. 4, 462-467.

Bolton, F. J., Sails, A. D., Fox, A. J., Wareing, D. R., and Greenway, D. L. 2002, Detection of *Campylobacter jejuni* and *Campylobacter coli* in foods by enrichment culture and polymerase chain reaction enzyme-linked immunosorbent assay., J. Food. Protect., vol. 65, no. 5, 760-767.

Bolton, F. J., Surman, S. B., Martin, K., Wareing, D. R., and Humphrey, T. J. 1999, Presence of *Campylobacter* and *Salmonella* in sand from bathing beaches, Epidemiol. Infect., vol. 122, no. 1, 7-13.

Boom, R., Sol, C. J. A., Salimans, M. M. M., Jansen, C. L. Werthen-Van Dillen P. M. E., and Van Der Noordaa, J. 1990, Rapid and Simple Method for Purification of Nucleic Acids, J. Clin. Microbiol., vol. 28, no. 3, 495-503.

Broman, T., Palmgren, H., Bergstrom, S., Sellin, M., Waldenstrom, J., Danielsson-Tham, M. L., and Olsen, B. 2002, *Campylobacter jejuni* in black-headed gulls (*Larus ridibundus*): prevalence, genotypes, and influence on *C. jejuni* epidemiology, J. Clin. Microbiol., vol. 40, no. 12, 4594-4602.

Brown, P., Kidd, D., Riordan, T., and Barrell, R. A. 1988, An outbreak of food-borne *Campylobacter jejuni* infection and the possible role of cross-contamination, J. Infect., vol. 17, no. 2, 171-176.

Burnett, T. A., Hornitzky, A., Kuhnert, P., and Djordjevic, S. P. 2002, Speciating *Campylobacter jejuni* and *Campylobacter coli* isolates from poultry and humans using six PCR-based assays, FEMS. Microbiol. Lett., vol. 216, no. 2, 201-209.

Butzler, J. P., Dekeyser, P., Detrain, M., and Dehaen, F. 1973, Related Vibrio in stools, J. Pediatrics, vol. 82, 493-495.

Butzler, J.P. 1982 Campylobacter Enteritis, Infection, vol, 10. suppl. 2, S67-69

Butzler, J. P. and Oosterom, J. 1991, Campylobacter:pathogenicity and significance in foods, Int. J. Food. Microbiol., vol. 12, 1-8.

Butzler, J. P. and Skirrow, M. B. 1979, Campylobacter enteritis, Clin. Gastro., vol. 8, 737-765.

Cappelier, J. M., Lazaro, B., Rossero, A., Fernandez-Astorga, A., and Federighi, M. 1997, Double staining (CTC-DAPI) for detection and enumeration of viable but non-culturable *Campylobacter jejuni* cells, Vet. Res., vol. 28, no. 6, 547-555.

Cappelier, J. M., Rossero, A., and Federighi, M. 2000, Demonstration of a protein synthesis in starved *Campylobacter jejuni* cells, Int. J. Food. Microbiol., vol. 55, no. 1-3, 63-67.

Cardarelli-Leite, P., Blom, K., Patton, C. M., Nicholson, M. A., Steigerwalt, A. G., Hunter S.B., Brenner, D. J., and Swaminathan, B. 1996, Rapid identification of *Campylobacter* species by restriction fragment length polymorphism analysis of a PCR-amplified fragment of the gene for 16S rRNA, J. Clin. Microbiol., vol. 34, no. 1, 62-67.

Carvalho, A. C., Ruiz-Palacios, G. M., Ramos-Cervantes, P., Cervantes, L. E., Jiang, X., and Pickering, L. K. 2001, Molecular characterization of invasive and noninvasive *Campylobacter jejuni* and *Campylobacter coli* isolates, J. Clin. Microbiol., vol. 39, no. 4, 1353-1359.

Casademont, I., Chevrier, D., and Guesdon, J. L. 1998, Cloning of a *sap*B homologue (*sap*B2) encoding a putative 112-kDa *Campylobacter fetus* S-layer protein and its use for identification and molecular genotyping, FEMS Immunol. Med. Microbiol., vol. 21, no. 4, 269-281.

Champion, O. L., Best, E. L., and Frost, J. A. 2002, Comparison of pulsed-field gel electrophoresis and amplified fragment length polymorphism techniques for investigating outbreaks of enteritis due to campylobacters, J. Clin. Microbiol., vol. 40, no. 6, 2263-2265.

Chen, S., Yee, A, Griffiths, M. Larkin C., Yamahiro, C. T., Behari, R., Paszko-Kolva, C., Rahn, K., and De Grandis, S. A. 1997, The Evaluation of a fluorogenic Polymerase Chain Reaction assay for the detection of *Salmonella* species in food commodities., Int. J. Food. Microbiol., vol. 35, 239-250.

Cheng, Z. and Griffiths, M. W. 2003, Rapid Detection of *Campylobacter jejuni* in Chicken rinse water by melting -peak analysis of amplicons in real-time polymerase chain reaction, J. Food. Prot., vol. 66, no. 8, 1343-1352.

Chuma, T., Hashimoto, S., and Okamoto, K. 2000, Detection of thermophilic campylobacter from sparrows by multiplex PCR: the role of sparrows as a source of contamination of broilers with Campylobacter, J. Vet. Med. Sci., vol. 62, no. 12, 1291-1295.

Clark, C. G., Price, L., Ahmed, R., Woodward, D. L., Melito, P. L., Rodgers, F. G., Jamieson, F., Ciebin, B., Li, A., and Ellis, A. 2003, Characterization of waterborne outbreak-associated *Campylobacter jejuni*, Walkerton, Ontario, Emerg. Infect. Dis., vol. 9, no. 10, 1232-1241.

Clarke, S. C. 2002, Nucleotide sequence-based typing of bacteria and the impact of automation, Bioessays, vol. 24, no. 9, 858-862.

Clarke, S. C., Diggle, P. J., and Edwards, G. F. S. 2001, Semiautomation of multilocus sequence typing for the characterisation of clinical isolates of *Neisseria meningitidis*, J. Clin. Microbiol., vol. 39, no. 9, 3066-3071.

Cloak, O. M., Duffy, G., Sheridan, J. J., Blair, I. S., and Mc Dowell, D. A. 2001, A survey on the incidence of *Campylobacter* spp. and the development of a surface adhesion polymerase chain reaction (SA-PCR) assay for the detection of *Campylobacter jejuni* in retail meat products., Food Microbiol., vol. 18, 287-298.

Cloak, O. M. and Fratamico, P. M. 2002, A multiplex polymerase chain reaction for the differentiation of *Campylobacter jejuni* and *Campylobacter coli* from a swine processing facility and characterization of isolates by pulsed-field gel electrophoresis and antibiotic resistance profiles, J. Food. Prot., vol. 65, no. 2, 266-273.

Coates, D., Hutchinson, D. N., and Bolton, F. J. 1987, Survival of thermophilic campylobacters on fingertips and their elimination by washing and disinfection, Epidemiol. Infect., vol. 99, no. 2, 265-274.

Colles, F. M., Jones, K., Harding, R. M., and Maiden, M. C. 2003, Genetic diversity of *Campylobacter jejuni* isolates from farm animals and the farm environment, Appl. Environ. Microbiol., vol. 69, no. 12, 7409-7413.

Cools, I., Uyttendaele, M., Caro, C., D'Haese, E., Nelis, H. J., and Debevere, J. 2003, Survival of *Campylobacter jejuni* strains of different origin in drinking water, J. Appl. Microbiol., vol. 94, no. 5, 886-892.

Corless, C. E., Guiver, m., Borrow, R., Edwards-Jones, V., and Fox, A. J. 2000, Contamination and sensitivity issues with real-time universal 16S rRNA PCR, J. Clin. Microbiol., vol. 38, no. 5, 1747-1752.

Corry, J. E., Post, D. E., Colin, P., and Laisney, M. J. 1995, Culture media for the isolation of campylobacters, Int. J. Food Microbiol., vol. 26, no. 1, 43-76.

Corry, J. E., Allen, P. M., Aydin, F., Atabay, H. I., Bull, S. A., Domingue, G., & Humphrey, T. H. 2001 Colonisation of broiler chickens with thermophillic campylobacters in three different

integrated poultry companies during rearing and slaughter, Int. J. Med. Microbiol., vol. 291 supp 31, 41.

Cowden, J. 1992, *Campylobacter*: epidemiological paradoxes, BMJ, vol. 305, no. 6846, 132-133.

D'Agata, E. M., Gerrits, M. M., Tang, Y. W., Samore, M., and Kusters, J. G. 2001, Comparison of pulsed-field gel electrophoresis and amplified fragment-length polymorphism for epidemiological investigations of common nosocomial pathogens, Infect. Control Hosp. Epidemiol., vol. 22, no. 9, 550-554.

Day, W. A., Pepper, I. L., and Joens, L. A. 1997, Use of a arbitrarily primed PCR product in the development of a *Campylobacter jejuni*-specific PCR, Appl. Environ. Microbiol., vol. 63, 1019-1023.

de Boer, P., Wagenaar, J. A., Achterberg, R. P., Putten, J. P., Schouls, L. M., and Duim, B. 2002, Generation of *Campylobacter jejuni* genetic diversity *in vivo*, Mol. Microbiol., vol. 44, no. 2, 351-359.

De Cesare, A., Sheldon, B. W., Smith, K. S., and Jaykus, L. A. 2003, Survival and persistence of *Campylobacter* and *Salmonella* species under various organic loads on food contact surfaces, J. Food Prot., vol. 66, no. 9, 1587-1594.

De Melo, M. A., Gabbiani, G., and Pechere, J. C. 1989, Cellular events and intracellular survival of *Campylobacter jejuni* during infection of HEp-2 cells, Infect. Immun., vol. 57, no. 7, 2214-2222.

Dekeyser, P., Gossuin-Detrain, J., Butzler, J. P., and Sternon, J. 1972, Acute enteritis due to related Vibrio: First positive stool cultures, J. Infect. Dis., vol. 125, 390-392.

Denis, M., Refregier-Petton, J., Laisney, M. J., Ermel, G., and Salvat, G. 2001, Campylobacter contamination in french chicken production from farm to consumers. Use of a PCR assay for detection and identification of *Campylobacter jejuni* and *Camp. coli*, J. Appl. Microbiol., vol. 91, no. 2, 255-267.

Denis, M., Soumet, C., Rivoal, K., Ermel, G., Salvat, G., and Colin.P. 1999, Development of a m-PCR Assay for simultaneous identification of *Campylobacter jejuni* and *C. coli*, Lett. Appl. Microbiol., vol. 29, 406-410.

Desai, M., Logan, J. M., Frost, J. A., and Stanley, J. 2001, Genome sequence-based fluorescent amplified fragment length polymorphism of *Campylobacter jejuni*, its relationship to serotyping, and its implications for epidemiological analysis, J. Clin. Microbiol., vol. 39, no. 11, 3823-3829.

Devost, N. C. and Choy, F. Y. 2000, Mutation analysis of Gaucher disease using dot-blood samples on FTA filter paper, Am. J. Med. Genet., vol. 94, no. 5, 417-420.

Dickins, M. A., Franklin, S., Stefanova, R., Schutze, G. E., Eisenach, K. D., Wesley, I, and Cave, M. D. 2002, Diversity of campylobacter isolates from retail poultry carcasses and from humans as demonstrated by pulsed-field gel electrophoresis, J. Food Prot., vol. 65, no. 6, 957-962.

Diggle, M. A. and Clarke, S. C. 2002, What a load of old sequence!!!, J. Clin. Microbiol., vol. 40, no. 7, 2707.

Dingle, K. E., Colles, F. M., Ure, R., Wagenaar, J. A., Duim, B., Bolton, F. J., Fox, A. J., Wareing, D. R., and Maiden, M. C. 2002, Molecular characterization of *Campylobacter jejuni* clones: a basis for epidemiologic investigation, Emerg. Infect. Dis., vol. 8, no. 9, 949-955.

Dingle, K. E., Colles, F. M., Wareing, D. R. A., Ure, R., Fox, A. J., Bolton, F. E., Bootsma, H. J., Willems, R. J. L., Urwin, R., and Maiden, M. C. J. 2001, Multilocus sequence typing system for *Campylobacter jejuni*, J. Clin. Microbiol., vol. 39, no. 1, 14-23.

Doig, P., Yao, R., Burr, D. H., Guerry, P., and Trust, T. J. 1996, An environmentally regulated pilus-like appendage involved in campylobacter pathogenesis, Mol. Microbiol., vol. 20, no. 4, 885-894.

Dorrell, N., Mangan, J.A., Laing, K.G., Hinds, J., Linton, D., Al-Ghusein, H., Barrell, B.G., Parkhill. J., Stoker, N.G., Karlyshev, A.V., Butcher, P.D. and Wren, B.W. 2001. Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. Genome Res. Vol. 11, 1706-15.

Doyle, M. P. and Roman, D. J. 1982, Response of *Campylobacter jejuni* to sodium chloride, Appl. Environ. Microbiol., vol. 43, 561-565.

Drosten, C., Seifried, E., and Roth, W. K. 2001, TaqMan 5'-nuclease human immunodeficiency virus type 1 PCR assay with phage-packaged competitive internal control for high-throughput blood donor screening, J. Clin. Microbiol., vol. 39, no. 12, 4302-4308.

Duim, B., Vandamme, P. A., Rigter, A., Laevens, S., Dijkstra, J. R., and Wagenaar, J. A. 2001, Differentiation of *Campylobacter* species by AFLP fingerprinting, Microbiology, vol. 147, no.10, 2729-2737.

Duim, B., Wassenaar, T.M., Rigter, A., Wagenaar, J. 1999, High-resolution genotyping of Campylobacter strains isolated from poultry and humans with amplified fragment length polymorphism fingerprinting, Appl. Environ. Microbiol. vol. 65, no. 6, 2369-2375.

Engberg, J., Gerner-Smidt, P., Scheutz, F., Moller, Nielsen E., On, S. L., and Molbak, K. 1998, Water-borne *Campylobacter jejuni* infection in a Danish town, a 6-week continuous source outbreak, Clin. Microbiol. Infect., vol. 4, no. 11, 648-656.

Englen, M. D. and Fedorka-Cray, P. J. 2002, Evaluation of a commercial diagnostic PCR for the identification of *Campylobacter jejuni* and *Campylobacter coli*, Lett. Appl. Microbiol., vol. 35, 353-356.

Englen, M. D., Ladely, S. R., and Fedorka-Cray, P. J. 2003, Isolation of campylobacter and identification by PCR, Methods Mol. Biol., vol. 216:109-21., 109-121.

Engvall, E. O., Brandstrom, B., Gunnarsson, A., Morner, T., Wahlstrom, H., and Fermer, C. 2002, Validation of a polymerase chain reaction/restriction enzyme analysis method for species identification of thermophilic campylobacters isolated from domestic and wild animals, J. Appl. Microbiol., vol. 92, no. 1, 47-54.

Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J., and Spratt, B. G. 2000, Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*, J. Clin. Microbiol., vol. 38, no. 3, 1008-1015.

Enright MC, Spratt BG. 1998 A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology. Vol. 144, no. 11, 3049-60.

Enright, M. C. and Spratt, B. G. 1999, Multilocus Sequence Typing, Trends Microbiol., vol. 7, no. 12, 482-487.

Enright, M. C., Spratt, B. G., Kalia, A., Cross, J. H., and Bessen, D. E. 2001, Multilocus sequence typing of *Streptococcus pyogenes* and the relationships between emm type and clone, Infect. Immun., vol. 69, no. 4, 2416-2427.

Ertas, H. B., Ozbey, G., Kilic, A., and Muz, A. 2003, Isolation of *Campylobacter jejuni* and *Campylobacter coli* from the gall bladder samples of sheep and identification by polymerase chain reaction, J. Vet. Med. B Infect. Dis. Vet. Public Health, vol. 50, no. 6, 294-297.

Escherich, T. 1886, Articles adding to the knowledge of intestinal bacteria, III On the existence of Vibrios in the intestines and feces of babies., Munch. Med. Wochenschr, vol. 33, 815-817.

Evans, M. R., Lane, W., Frost, J. A., and Nylen, G. 1998, A campylobacter outbreak associated with stir fried food, Epidemiol. Infect., vol. 121, 275-279.

Evans, M. R., Ribeiro, C. D., and Salmon, R. L. 2003, Hazards of healthy living: bottled water and salad vegetables as risk factors for campylobacter infection, Emerg. Infect. Dis., vol. 9, no. 10, 1219-1225.

Evans, S. J. and Sayers, A. R. 2000, A longitudinal study of campylobacter infection of broiler flocks in Great Britain, Prev. Vet. Med., vol. 46, no. 3, 209-223.

Everest, P. H., Goossens, H., Butzler, J. P., Lloyd, D., Knutton, S., Ketley, J. M., and Williams, P. H. 1992, Differentiated Caco-2 cells as a model for enteric invasion by *Campylobacter jejuni* and *C. coli*, J. Med. Microbiol., vol. 37, no. 5, 319-325.

Everest, P. H., Goossens, H., Sibbons, P., Lloyd, D. R., Knutton, S., Leece, R., Ketley, J. M., and Williams, P. H. 1993, Pathological changes in the rabbit ileal loop model caused by *Campylobacter jejuni* from human colitis, J. Med. Microbiol., vol. 38, no. 5, 316-321.

Eyers, M., Chapelle, S., Van Camp, G., Goossens, H., and De Watcher, R. 1993, Discrimination among thermophillic *Campylobacter* species by polymerase chain reaction amplification of 23S rRNA gene fragments, J. Clin. Microbiol., vol. 31, 3340-3343.

Feng, P. 1997, Impact of molecular biology on the detection of food-borne pathogens, Mol. Biotech., vol. 7, 267-278.

Fermandez, H., Vergara, M., and Tapia, F. 1985, Dessication resistance in thermotolerant *Campylobacter* species, Infection, vol. 13, 197.

Field, L. H., Headley, V. L., Payne, S. M., and Berry, L. J. 1986, Influence of iron on growth, morphology, outer membrane protein composition, and synthesis of siderophores in *Campylobacter jejuni*, Infect. Immun., vol. 54, no. 1, 126-132.

Florent, A. 1959, "Two Forms of Bovine Genital Vibriosis, Venereal Vibrosis due to *V.foetus venerialis* and Vibriosis of Intestinal origin due to *V.foetus intestinalis.*", Meded. Veeartsenijisch. Rijksuniv. Gent., vol. 3. 953-957.

Florin, I. and Antillon, F. 1992, Production of enterotoxin and cytotoxin in *Campylobacter jejuni* strains isolated in Costa Rica, J. Med. Microbiol., vol. 37, no. 1, 22-29.

Freidman, C. R., Neiman, J., Wegener, H. C., & Tauxe, R. V. 2000, "Epidemiology of Campylobacter Infections in the United States and other Industrialised Nations," in *Campylobacter*, 2 edn, I. Nachampkin & J. M. Blaser, eds., American Society for Microbiology, Washington D.C., pp. 126-127.

Frost, J. A. 2001, Current Epidemiological Issues in Human Campylobacteriosis, J. Appl. Microbiol, vol. 90, 85s-95s.

Frost, J. A., Gillespie I.A., and O'Brien, S. J. 2002, Public health implications of campylobacter outbreaks in England and Wales, 1995-9: epidemiological and microbiological investigations, Epidemiol. Infect, vol. 128, 111-118.

Frost, J. A., Kramer, J. M., and Gillanders, S. A. 1999, Phage typing of *C.jejuni* and *C.coli* and its use as an adjunct to serotyping, Epidemiol. Infect., vol. 123, 47-55.

Frost, J. A., Oza, A. N., and Thwaites, R. T. 1998, Serotyping scheme for *C.jejuni* and *C.coli* based on direct agglutination of heat stable antigens., J. Clin. Microbiol., vol. 36, 335-339.

Fujita, M., Fujimoto, S., Morooka, T., and Amako, K. 1995, Analysis of strains of *Campylobacter fetus* by pulsed-field gel electrophoresis, J. Clin. Microbiol., vol. 33, no. 6, 1676-1678.

Gaynor, E. C., Cawthraw, S., Manning, G., MacKichan, J. K., Falkow, S., and Newell, D. G. 2004, The genome-sequenced variant of *Campylobacter jejuni* NCTC 11168 and the original clonal clinical isolate differ markedly in colonization, gene expression, and virulence-associated phenotypes, J. Bacteriol., vol. 186, no. 2, 503-517.

Gaynor, E. C., Ghori, N. and Falkow, S. 2001, Bile-induced "pili" in *Campylobacter jejuni* are bacteria-independent artefacts of the culture medium. Mol. Microbiol., vol. 39, no. 6, 1546-1549.

Ge, B., Zhao, S., Hall, R., and Meng, J. 2002, A PCR-ELISA for detecting Shiga toxin-producing *Escherichia coli*, Microbes. Infect., vol. 4, no. 3, 285-290.

Germer, J. J., Lins, M. M., Jensen, M. E., Harmsen, W. S., Ilstrup, D. M., Mitchell, P. S., Cockerill, F. R., HI, and Patel, R. 2003, Evaluation of the MagNA pure LC instrument for extraction of hepatitis C virus RNA for the COBAS AMPLICOR Hepatitis C Virus Test, version 2.0, J. Clin. Microbiol., vol. 41, no. 8, 3503-3508.

Gibbens, J. C., Pascoe, S. J., Evans, S. J., Davies, R. H., and Sayers, A. R. 2001, A trial of biosecurity as a means to control campylobacter infection of broiler chickens, Prev. Vet. Med, vol. 48, no. 2, 85-99.

Gibson, J. R., Fitzgerald, C., and Owen, R. J. 1995, Comparison of PFGE, Ribotyping and Phage typing in Epidemiological Analysis of *Campylobacter jejuni* Serotype HS2 infections., Epidemiol. Infect, vol. 115, 215-225.

Gibson, J. R., Slater, E., Xerry, J., Tompkins, D. S., and Owen, R. J. 1998, Use of an amplified-fragment length polymorphism technique to fingerprint and differentiate isolates of *Helicobacter pylori*, J. Clin. Microbiol., vol. 36, no. 9, 2580-2585.

Gillespie, I. A., Adak, G. K., O'Brien, S. J., and Bolton, F. J. 2003, Milkborne general outbreaks of infectious intestinal disease, England and Wales, 1992-2000, Epidemiol. Infect., vol. 130, no. 3, 461-468.

Gillespie, I. A., O'Brien, S. J., and Frost, J. A. 2001, Campylobacter Sentinel Surveillance Scheme, Results from the First Year of Study., Not in press.

Gillespie, I. A., O'Brien, S. J., Frost, J. A., Adak, G. K., Horby, P., Swan, A. V., Painter, M. J., Neal, K. R., Campylobacter Sentinel, Surveillance Scheme, and Collaborators. 2002, A case-case comparison of *Campylobacter coli* and *Campylobacter jejuni* infection: a tool for generating hypotheses, Emerg. Infect. Dis., vol. 8, no. 9, 937-942.

Gisendorf, G. A. J., Quint, W. G. V., Henkens, M. H. C., Stegeman, H., Huf, F. A., and Niesters, H. G. M. 1992, Rapid and sensitive detection of *Campylobacter* spp. in chicken products by using the polymerase chain reaction, Appl. Environ. Microbiol., vol. 58, no. 12, 3804-3808.

Gonzalez, I., Grant, K. A., Richardson, P. T., and Park, S. F. Collins 1997, Specific identification of the enteropathogens *Campylobacter jejuni* and *Campylobacter coli* by using a PCR test based on the *CeuE* gene encoding a putative virulence determinant., J. Clin. Microbiol., vol. 35, no. 3, 759-763.

Goossens, H., Rummens, E., Cadranel, S., Butzler, J. P., and Takeda, Y. 1985, Cytotoxic activity on chinese hamster ovary cells in culture filtrates of *Campylobacter jejuni/coli*, Lancet, vol. 2, no. 8453, 511.

Gorkiewicz, G., Feierl, G., Schober, C., Dieber, F., Kofer, J., Zechner, R., and Zechner, E. L. 2003, Species-specific identification of campylobacters by partial 16S rRNA gene sequencing, J. Clin. Microbiol., vol. 41, no. 6, 2537-2546.

Gorkiewicz, G., Feierl, G., Zechner, R., and Zechner, E. L. 2002, Transmission of *Campylobacter hyointestinalis* from a pig to a human, J. Clin. Microbiol., vol. 40, no. 7, 2601-2605.

Grajewski, B. A., Kusek, J. W., and Gelfand, H. M. 1985, Development of a bacteriophage typing scheme for *Campylobacter jejuni* and *Campylobacter coli*, J. Clin. Microbiol., vol. 22, 13-18.

Grant, C. C., Konkel, M. E., Cieplak, W., Jr., and Tompkins, L. S. 1993, Role of flagella in adherence, internalization, and translocation of *Campylobacter jejuni* in nonpolarized and polarized epithelial cell cultures, Infect.Immun., vol. 61, no. 5, 1764-1771.

Gregory, E., Barnhart, H., Dreesen, D. W., Stern, N. J., and Corn, J. L. 1997, Epidemiological study of *Campylobacter spp.* in broilers: source, time of colonization, and prevalence, Avian Dis., vol. 41, no. 4, 890-898.

Guerry, P., Perez-Casal, J., Yao, R., McVeigh, A., and Trust, T. J. 1997, A genetic locus involved in iron utilization unique to some campylobacter strains, J. Bacteriol., vol. 179, no. 12, 3997-4002.

Guiver, M., Borrow, R., Marsh, J., Gray, S. J., Kaczmarski, E. B., Howells, D., Boseley, P., and Fox, A. J. 2000, Evaluation of the Applied Biosystems automated Taqman polymerase chain reaction system for the detection of meningococcal DNA, FEMS Immunol. Med. Microbiol., vol. 28, no. 2, 173-179.

Gutacker, M. M., Smoot, J. C., Migliaccio, C. A., Ricklefs, S. M., Hua, S., Cousins, D. V., Graviss, E. A., Shashkina, E., Kreiswirth, B. N., and Musser, J. M. 2002, Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains, Genetics, vol. 162, no. 4, 1533-1543.

Hani, E. K. and Chan, V. L. 1995, Expression and characterization of *Campylobacter jejuni* benzoylglycine amidohydrolase (Hippuricase) gene in *Escherichia coli*, J. Bacteriol., vol. 177, no. 9, 2396-2402.

Hanninen, M., Hakkinen, M., and Rautelin, H. 1999, Stability of related human and chicken *Campylobacter jejuni* genotypes after passage through chick intestine studied by Pulsed Field Gel Electrophoresis, Appl. Environ. Microbiol., vol. 65, no. 5, 2272-2275.

Hanninen, M. L, Perko-Makela, P., Rautelin, H., Dium, B., and Waggenaar, J. A. 2001, Genomic relatedness within five finnish *Campylobacter jejuni* pulsed-field gel electrohoresis genotypes studied by amplified fragment length polymorphism analysis, ribotyping and serotyping., Appl. Environ. Microbiol., vol. 67, no. 4, 1581-1585.

Hanninen, M. L., Sarelli, L., Sukura, A., On, S. L., Harrington, C. S., Matero, P., and Hirvela-Koski, V. 2002, *Campylobacter hyointestinalis* subsp. *hyointestinalis*, a common *Campylobacter* species in reindeer, J. Appl. Microbiol., vol. 92, no. 4, 717-723.

Hansen, P. and Blakesley, R. 1998, Simple archiving of bacterial and plasmid DNAs for future use, Focus, vol. 20, no. 3, 72-74.

Harmon, K. M., Ramson, G. M., and Wesley, I. R. 1997, Differentiation of *Campylobacter jejuni* and *Campyobacter coli* by polymerase chain reaction, Mol. Cell. Probes, vol. 11, 195-200.

Harrington, C. S., Thomson-Carter, F., and Carter, P. E. 1999, Molecular epidemiological investigation of an outbreak of *Campylobacter jejuni* identifies a dominant clonal line within Scottish serotype HS55 populations, J. Clin. Microbiol., vol. 41, 15-26.

Harris, N. V., Weiss, N. S., and Nolan, C. M. 1986, The role of poultry and meats in the etiology of *Campylobacter jejuni/coli* enteritis, Am. J. Public Health, vol. 76, no. 4, 407-411.

Hein, I., Schneck, C., Knogler, M., Feierl, G., Plesss, P., Kofer, J., Achmann, R., and Wagner, M. 2003, *Campylobacter jejuni* isolated from poultry and humans in Styria, Austria: epidemiology and ciprofloxacin resistance, Epidemiol. Infect., vol. 130, no. 3, 377-386.

Hickey, T. E., McVeigh, A. L., Scott, D. A., Michielutti, R. E., Bixby, A., Carroll, S. A., Bourgeois, A. L., and Guerry, P. 2000, *Campylobacter jejuni* cytolethal distending toxin mediates release of interleukin-8 from intestinal epithelial cells, Infect. Immun., vol. 68, no. 12, 6535-6541.

Higgins, J. A., Jenkins, M. C., Shelton, D. R., Fayer, R., and Karns, J. S. 2001, Rapid extraction of DNA From *Escherichia coli* and *Cryptosporidium parvum* for use in PCR, Appl. Environ. Microbiol., vol. 67, no. 11, 5321-5324.

Hodgson, A. E., McBride, B. W., Hudson, M. J., Hall, G., and Leach, S. A. 1998, Experimental campylobacter infection and diarrhoea in immunodeficient mice, J. Med. Microbiol., vol. 47, no. 9, 799-809.

Hoffmaster, A. R., Fitzgerald, C. C., Ribot, E., Mayer, L. W., and Popovic, T. 2002, Molecular subtyping of *Bacillus anthracis* and the 2001 bioterrorism-associated anthrax outbreak, United States, Emerg. Infect. Dis., vol. 8, no. 10, 1111-1116.

Holland, P. M., Abramson, R. D., Watson, R., and Gelfand, D. H. 1991, Detection of specific polymerase chain reaction product by utilizing the 5'→3' exonuclease activity of *Thermus aquaticus* DNA polymerase, Proc. Natl. Acad. Sci U.S.A, vol. 88, no. 16, 7276-7280.

Homan, W. L., Tribe, D., Poznanski, S., Li, M., Hogg, G., Spalburg, E., Van Embden, J. D., and Willems, R. J. 2002, Multilocus sequence typing scheme for *Enterococcus faecium*, J. Clin. Microbiol., vol. 40, no. 6, 1963-1971.

Hossain, A., Stewart-Tull, D. E., and Freer, J. H. 1993, Heat-labile and heat-stable haemolysins of *Campylobacter jejuni*, FEMS Immunol. Med. Microbiol., vol. 6, no. 4, 331-339.

Hu, L. & Kopecko, D. J. 2002, "Interactions of Campylobacter cells:Gut luminal colonisation and mucosal invasion mechanisms," in *Campylobacter*, I. Nachamkin & J. M. Blaser, eds., American Society for Microbiology, Washington D.C., pp. 191-215.

Hudson, S. J., Lightfoot, N. F., Coulson, J. C., Russell, K., Sisson, P. R., and Sobo, A. O. 1991, Jackdaws and magpies as vectors of milkborne human campylobacter infection, Epidemiol. Infect., vol. 107, no. 2, 363-372.

Hugenholtzt, P. and Huber, T. 2003, Chimeric 16S rDNA sequences of diverse origin are accumulating in the public databases, Int. J. Syst. Evol. Microbiol., vol. 53, no. 1, 289-293.

Hume, M. E., Droleskey, R. E., Sheffield, C. L., and Harvey, R. B. 2002, *Campylobacter coli* pulsed field gel electrophoresis genotypic diversity among sows and piglets in a farrowing barn, Curr. Microbiol., vol. 45, no. 2, 128-132.

Humphrey, T. 2001, The spread and persistence of *Campylobacter* and *Salmonella* in the domestic kitchen, J. Infect., vol. 43, no. 1, 50-53.

Humphrey, T. J., Martin, K. W., Slader, J., and Durham, K. 2001, *Campylobacter* spp. in the kitchen: spread and persistence, Symp. Ser. Soc. Appl. Microbiol. no. 30, 115S-120S.

Hurtado, A. and Owen, R. J. 1997, A Molecular scheme based on 23S rRNA gene polymorphisms for rapid identification of *Campylobacter* and *Arcobacter* species, J. Clin. Microbiol., vol. 35, no. 9, 2401-2404.

Hutchinson, D. N. and Bolton, F. J. 1983, Is enrichment culture necessary for the isolation of *Campylobacter jejuni* from faeces?, J. Clin. Pathol, vol. 36, no. 12, 1350-1352.

Hutchinson, D. N. and Bolton, F. J. 1984, Improved blood free selective medium for the isolation of *Campylobacter jejuni* from faecal specimens, J. Clin. Pathol., vol. 37, no. 8, 956-957.

Hutchinson, D. N., Bolton, F. J., Jelley, W. C., Mathews, W. G., Telford, D. R., Counter, D. E., Jessop, E. G., and Horsley, S. D. 1985, Campylobacter enteritis associated with consumption of raw goat's milk, Lancet, vol. 1, no. 8436, 1037-1038.

Inglis, G. D. and Kalischuk, L. D. 2003, Use of PCR for direct detection of *Campylobacter* species in bovine feces, Appl. Environ. Microbiol., vol. 69, no. 6, 3435-3447.

Iredell, J., Blanckenberg, D., Arvand, M., Grauling, S., Feil, E. J., and Birtles, R. J. 2003, Characterization of the natural population of *Bartonella henselae* by Multilocus Sequence Typing, J. Clin. Microbiol., vol. 41, no. 11, 5071-5079.

Jackson, C. J., Fox, A. J., and Jones, D. M. 1996, A novel polymerase chain reaction assay for the detection and speciation of thermophilic *Campylobacter* spp., J. Appl. Microbiol., vol. 81, 467-473.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M., and Kersters, K. 1996, Evaluation of the DNA fingerprinting method AFLP as an new tool in bacterial taxonomy, Microbiology, vol. 142, no. 7, 1881-1893.

Jolley, K. A., Kalmusova, J., Feil, E. J., Gupta, S., Musilek, M., Kriz, P., and Maiden, M. C. 2000, Carried meningococci in the Czech Republic: a diverse recombining population, J. Clin. Microbiol., vol. 38, no. 12, 4492-4498.

Jones, D. M., Sutcliffe, E. M., Rios, R., Fox, A. J., and Curry, A. 1993, *Campylobacter jejuni* adapts to aerobic metabolism in the environment, J. Med. Microbiol., vol. 38, no. 2, 145-150.

Jones, F. S., Orcutt, M., and Little, R. B. 1931, Vibrios associated with intestinal disorders of cows and calves, J. Exp. Med., vol. 53, no. 853, 864.

Jones, K. 2001, The *Campylobacter* conundrum, Trends Microbiol., vol. 9, no. 8, 365-366.

Jones, N., Bohnsack, J. F., Takahashi, S., Oliver, K. A., Chan, M. S., Kunst, F., Glaser, P., Rusniok, C., Crook, D. W., Harding, R. M., Bisharat, N., and Spratt, B. G. 2003, Multilocus sequence typing system for group B streptococcus, J. Clin. Microbiol., vol. 41, no. 6, 2530-2536.

Jorgensen, F., Bailey, R., Williams, S., Henderson, P., Wareing, D. R., Bolton, F. J., Frost, J. A., Ward, L., and Humphrey, T. J. 2002, Prevalence and numbers of *Salmonella* and *Campylobacter* spp. on raw, whole chickens in relation to sampling methods, Int. J. Food Microbiol., vol. 76, no. 1-2, 151-164.

Jurevic, R. J., Bai, M., Chadwick, R. B., White, T. C., and Dale, B. A. 2003, Single-nucleotide polymorphisms (SNPs) in human beta-defensin 1: high-throughput SNP assays and association with *Candida* carriage in type I diabetics and nondiabetic controls, J. Clin. Microbiol., vol. 41, no. 1, 90-96.

Kaldor, J. and Speed, B. R. 1984, Guillain-Barre syndrome and *Campylobacter jejuni*: a serological study, Br. Med. J.(Clin.Res.Ed), vol. 288, no. 6434, 1867-1870.

Kalinina, O., Lebedeva, I., Brown, J., and Silver, J. 1997, Nanoliter scale PCR with TaqMan detection, Nucleic Acids Res., vol. 25, no. 10, 1999-2004.

Kalman, M., Szollosi, E., Czermann, B., Zimanyi, M., Szekeres, S., and Kalman, M. 2000, Milkborne campylobacter infection in Hungary, J. Food Prot., vol. 63, no. 10, 1426-1429.

Karlyshev, A. V., McCrossan, M. V., and Wren, B. W. 2001, Demonstration of polysaccharide capsule in *Campylobacter jejuni* using electron microscopy, Infect. Immun., vol. 69, no. 9, 5921-5924.

Keat, A. and Rowe, I. 1991, Reiter's syndrome and associated arthritides, Rheum. Dis. Clin. North Am., vol. 17, no. 1, 25-42.

Kelly, D. J. 2001, The physiology and metabolism of *Campylobacter jejuni* and *Helicobacter pylori*, Symp. Ser. Soc. Appl. Microbiol. no. 30, 16S-24S.

Kendall, E. J. C. and Tanner, E. I. 1982, Campylobacter enteritis in general practice, J. Hygiene, vol. 88, 155-163.

Ketley, J. M. 1997, Pathogenesis of enteric infection by campylobacter, Microbiology, vol. 143, no. 1, 5-21.

Khakhira, R. and Lior, H. 1992, Extended phage-typing scheme for *Campylobacter jejuni* and *Campylobacter coli*, Epidemiol. Infect., vol. 108, 403-414.

Kim, N. W., Bingham, R., Louie, H., Hani, K., and Chan, V. L. 1992, Physical map of *Campylobacter jejuni* TGH9011 and localisation of 10 genetic markers by the use of Pulsed-Field Gel Electrophoresis, J. Bacteriol., vol. 174, no. 11, 3494-3498.

Kimura, M. 1991, The neutral theory of molecular evolution: a review of recent evidence, Jpn. J. Genet., vol. 66, no. 4, 367-386.

King, E. O. 1957, Human Infections with *Vibrio fetus* and a closely related vibrio, J. Infect. Dis., vol. 101, 119-128.

King, E. O. 1962, The laboratory recognition of *Vibrio fetus* and a closely related Vibrio from cases of human Vibriosis, Ann. N.Y. Acad. Sci., no. 98, 700-711.

Kirst, M. 1985, "The historical background of *Campylobacter* infection: new aspects.," in *Proccedings of the 3rd International Workshop on Campylobacter Infections;* A. D. Pearson, ed., Ottawa, pp. 23-27.

Kita, E., Nishikawa, F., Kamikaidou, N., Nakano, A., Katsui, N., and Kashiba, S. 1992, Mononuclear cell response in the liver of mice infected with hepatotoxigenic *Campylobacter jejuni*, J. Med. Microbiol., vol. 37, no. 5, 326-331.

Kita, E., Oku, D., Hamuro, A., Nishikawa, F., Emoto, M., Yagyu, Y., Katsui, N., and Kashiba, S. 1990, Hepatotoxic activity of *Campylobacter jejuni*, J. Med. Microbiol., vol. 33, no. 3, 171-182.

Klipstein, F. A., Engert, R. F., Short, H., and Schenk, E. A. 1985, Pathogenic properties of *Campylobacter jejuni*: assay and correlation with clinical manifestations, Infect. Immun., vol. 50, no. 1, 43-49.

Knepp, J. H., Geahr, M. A., Forman, M. S., and Valsamakis, A. 2003, Comparison of automated and manual nucleic acid extraction methods for detection of enterovirus RNA, J. Clin. Microbiol., vol. 41, no. 8, 3532-3536.

Konkel, M. E., Gray, S. A., Kim, B. J., Garvis, S. G., and Yoon, J. 1999, Identification of the enteropathogens *Campylobacter jejuni* and *Campylobacter coli* based on the *cadF* virulence gene and its product, J. Clin. Microbiol., vol. 37, 510-517.

Konkel, M. E. and Joens, L. A. 1989, Adhesion to and invasion of HEp-2 cells by *Campylobacter* spp, Infect. Immun., vol. 57, no. 10, 2984-2990.

Koo, K. and Jaykus, L. A. 2002, Detection of single nucleotide polymorphisms within the *Listeria* genus using an 'asymmetric' fluorogenic probe set and fluorescence resonance energy transfer based-PCR, Lett. Appl. Microbiol., vol. 35, no. 6, 513-517.

Koort, J. M., Lukinmaa, S., Rantala, M., Unkila, E., and Siitonen, A. 2002, Technical improvement to prevent DNA degradation of enteric pathogens in pulsed-field gel electrophoresis, J. Clin. Microbiol., vol. 40, no. 9, 3497-3498.

Kopecko, D. J., Hu, L., and Zaal, K. J. 2001, *Campylobacter jejuni,* microtubule-dependent invasion, Trends Microbiol., vol. 9, no. 8, 389-396.

Kotetishvili, M., Stine, O. C., Chen, Y., Kreger, A., Sulakvelidze, A., Sozhamannan, S., and Morris, J. G., Jr. 2003, Multilocus sequence typing has better discriminatory ability for typing *Vibrio cholerae* than does pulsed-field gel electrophoresis and provides a measure of phylogenetic relatedness, J. Clin. Microbiol., vol. 41, no. 5, 2191-2196.

Kramer, J. M., Frost, J. A., Bolton, F. J., and Wareing, D. R. A. 2000, Campylobacter contamination of raw meat and poultry at retail sale: identification of multiple types and comparison with isolates from human infection., J. Food. Prot., vol. 63, no. 12, 1654-1659.

Krause, R., Ramschak-Schwarzer, S., Gorkiewicz, G., Schnedl, W. J., Feierl, G., Wenisch, C., and Reisinger, E. C. 2002, Recurrent septicemia due to *Campylobacter fetus* and *Campylobacter lari* in an immunocompetent patient. Infection, vol. 30, no. 3, 171-174.

Kulkarni, S. P., Lever, S., Logan, J. M., Lawson, A. J., Stanley, J., and Shafi, M. S. 2002, Detection of *Campylobacter* species: a comparison of culture and polymerase chain reaction based methods, J. Clin. Path., vol. 55, no. 10, 749-753.

Kuroki, S., Saida, T., Nukina, M., Haruta, T., Yoshioka, M., Kobayashi, Y., and Nakanishi, H. 1993, *Campylobacter jejuni* strains from patients with Guillain-Barre syndrome belong mostly to Penner serogroup 19 and contain beta-N-acetylglucosamine residues, Ann. Neurol., vol. 33, no. 3, 243-247.

Lambert, M. A., Patton, C. M., Barrett, T. J., and Moss, C. W. 1987, Differentiation of *Campylobacter* and *Campylobacter*-like organisms by cellular fatty acid composition, J. Clin. Microbiol., vol. 25, no. 4, 706-713.

Lampel, K. A., Orlandi, P. A., and Kornegay, L. 2000, Improved template preparation for PCR-based assays for detection of food-borne bacterial pathogens, Appl. Environ. Microbiol., vol. 66, no. 10, 4539-4542.

Lander, K. P. and Gill, K. P. 1980, Experimental infection of the bovine udder with *Campylobacter coli/jejuni,* J. Hyg. (Lond), vol. 84, no. 3, 421-428.

Lawson, A. J., Linton, D., Stanley, J., and Owen, R. J. 1997, Polymerase chain reaction and speciation of *Campylobacter upsaliensis* and *C. helveticus* in human faeces and comparison with culture techniques., J. Appl. Microbiol., vol. 83, 375-380.

Lawson, A. J., Logan, J. M., O'neill, G. L., Desai, M., and Stanley, J. 1999, Large-scale survey of *Campylobacter* species in human gastroenteritis by PCR and PCR-enzyme-linked immunosorbent assay, J. Clin. Microbiol., vol. 37, no. 12, 3860-3864.

Lawson, A. J., Shafi, M. S., Pathak, K., and Stanley, J. 1998, Detection of campylobacter in Gastroenteritis: comparison of direct PCR assay of faecal samples with selective culture., Epidemiol. Infect., vol. 121, 547-553.

Lecuit, M., Abachin, E., Martin, A., Poyart, C., Pochart, P., Suarez, F., Bengoufa, D., Feuillard, J., Lavergne, A., Gordon, J. I., Berche, P., Guillevin, L., and Lortholary, O. 2004, Immunoproliferative small intestinal disease associated with *Campylobacter jejuni,* N. Engl. J. Med., vol. 350, no. 3, 239-248.

Lehner, A., Schneck, C., Feirl, G., Pless, P., Deutz, A., Brandl, E., and Wagner, M. 2000, Epidemiologic application of Pulsed-Field gel Electrophoresis to an outbreak of *Campylobacter jejuni* in an Austrian youth centre, Epidemiol. Infect., vol. 125, 13-16.

Levy, A. J. 1946, A gastro enteritis outbreak probably due to a bovine strain of Vibrio., Yale J. Biol. Med., vol. 18, 243-258.

Liesegang, A. and Tschape, H. 2002, Modified pulsed-field gel electrophoresis method for DNA degradation-sensitive *Salmonella enterica* and *Escherichia coli* strains, Int. J. Med. Microbiol., vol. 291, no. 8, 645-648.

Lindqvist, R., Norling, B., and Lambertz, S. T. 1997, A rapid sample preparation method for PCR detection of food pathogens based on buoyant density centrifugation, Lett. Appl. Microbiol., vol. 24, no. 4, 306-310.

Lindstedt, B. A., Heir, E., Gjernes, E., Vardund, T., and Kapperud, G. 2003, DNA fingerprinting of Shiga-toxin producing *Escherichia coli* O157 based on Multiple-Locus Variable-Number Tandem-Repeats Analysis (MLVA), Ann. Clin. Microbiol. Antimicrob., vol. 2, no. 1, 12.

Linton, D., Dewhirst, F. E., Clewley, J. P., Owen, R. J., Burnens, A. P., and Stanley, J. 1994, Two types of 16S rRNA gene are found in *Campylobacter helveticus*: analysis, applications and characterization of the intervening sequence found in some strains, Microbiology, vol. 140, no. 4, 847-855.

Linton, D., Lawson, A. J., Owen, R. J., and Stanley, J. 1997, PCR detection, identification to species level, and fingerprinting of *Campylobacter jejuni* and *Campylobacter coli* direct from diarrheic samples, J.Clin.Microbiol., vol. 35, no. 10, 2568-2572.

Linton, D., Owen, R. J., and Stanley, J. 1996, Rapid identification by PCR of the genus *Campylobacter* and of five *Campylobacter* species enteropathogenic for man and animals, Res. Microbiol., vol. 147, no. 9, 707-718.

Lior, H. 1982, Serotyping of *Campylobacter jejuni* by slide agglutination based on heat labile antigenic factors, J. Clin. Microbiol., vol. 36, 335-339.

Liu, Y. and Li, Y. 2002, Detection of *Escherichia coli* O157:H7 using immunomagnetic separation and absorbance measurement, J. Microbiol. Methods, vol. 51, no. 3, 369-377.

Llovo, J., Mateo, E., Munoz, A., Urquijo, M., On, S. L., and Fernandez-Astorga, A. 2003, Molecular typing of *Campylobacter jejuni* isolates involved in a neonatal outbreak indicates nosocomial transmission, J. Clin. Microbiol., vol. 41, no. 8, 3926-3928.

Logan, J. M. 2001, Rapid identification of *Campylobacter* spp. by melting peak analysis of biprobes in real-time PCR, J. Clin. Microbiol., vol. 39 No 6, 2227-2232.

Logan, J. M., Burnens, A., Linton, D., Lawson, A. J., and Stanley, J. 2000, *Campylobacter lanienae* sp. nov., a new species isolated from workers in an abattoir, Int. J. Syst. Evol. Microbiol., vol. 50 Pt 2:865-72., 865-872.

Lohmann, S., Lehmann, L., and Tabiti, K. 2000, Fast and flexible single nucleotide polymorphism (SNP) detection with the Lightcycler system, Biochemica, vol. 4, 23-28.

Lorenz, E., Lastovica, A., and Owen, R. J. 1998, Penner Serotypes 9,38 and 63, from human infections, animals and water by PFGE and flagellin gene analysis, Lett. Appl. Microbiol., vol. 26, 179-182.

Lucore, A. L., Cullison, M. A., and Jaykus, L. A. 2000, Immobilization with metal hydroxides as a means to concentrate food-borne bacteria for detection by cultural and molecular methods., Appl. Environ. Microbiol., vol. 66, no. 5, 1769-1776.

Magistrado, P. A., Garcia, M. M., and Raymundo, A. K. 2001, Isolation and polymerase chain reaction-based detection of *Campylobacter jejuni* and *Campylobacter coli* from poultry in the Philippines, Int. J. Food Microbiol., vol. 70, no. 1-2, 197-206.

Maher, M., Finnegan, C., Collins, E., Ward, B., Carroll, C., and Cormican, M. 2003, Evaluation of culture methods and a DNA probe-based PCR assay for detection of *Campylobacter* species in clinical specimens of feces, J. Clin. Microbiol., vol. 41, no. 7, 2980-2986.

Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. 1998, Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms, Proc. Natl. Acad. Sci U.S.A, vol. 95, no. 6, 3140-3145.

Manfreda, G., De Cesare, A., Bondioli, V., and Franchini, A. 2003a, Comparison of the BAX System with a multiplex PCR method for simultaneous detection and identification of *Campylobacter jejuni* and *Campylobacter coli* in environmental samples, Int. J. Food Microbiol., vol. 87, no. 3, 271-278.

Manfreda, G., De Cesare, A., Bondioli, V., and Franchini, A. 2003b, Ribotyping characterisation of campylobacter isolates randomly collected from different sources in Italy, Diagn. Microbiol. Infect. Dis., vol. 47, no. 2, 385-392.

Manning, G., Dowson, C. G., Bagnall, M. C., Ahmed, I. H., West, M., and Newell, D. G. 2003, Multilocus sequence typing for comparison of veterinary and human isolates of *Campylobacter jejuni*, Appl. Environ. Microbiol., vol. 69, no. 11, 6370-6379.

Matsheka, M. I., Elisha, B. G., Lastovica, A. L., and On, S. L. 2002, Genetic heterogeneity of *Campylobacter concisus* determined by pulsed field gel electrophoresis-based macrorestriction profiling, FEMS Microbiol.Lett., vol. 211, no. 1, 17-22.

Matsuda, M., Kaneko, A., Stanley, T., Millar, B. C., Miyajima, M., Murphy, P. G., and Moore, J. E. 2003, Characterization of urease-positive thermophilic *Campylobacter* subspecies by multilocus enzyme electrophoresis typing, Appl. Environ. Microbiol., vol. 69, no. 6, 3308-3310.

Matsuda, M., Shibuya, T., Itoh, Y., Takiguchi, M., Furuhata, K., Moore, J. E., Murayama, O., and Fukuyama, M. 2002, First isolation of urease-positive thermophilic Campylobacter (UPTC) from crows (*Corvus levaillantii*) in Japan, Int. J. Hyg. Environ. Health, vol. 205, no. 4, 321-324.

Mattick, K., Durham, K., Domingue, G., Jorgensen, F., Sen, M., Schaffner, D. W., and Humphrey, T. 2003, The survival of foodborne pathogens during domestic washing-up and subsequent transfer onto washing-up sponges, kitchen surfaces and food, Int. J. Food Microbiol., vol. 85, no. 3, 213-226.

McCardell, B. A., Madden, J. M., and Stanfield, J. T. 1986, Production of cytotoxins by Campylobacter, Lancet, vol. 1, no. 8488, 1031.

McFadyean, J. & Stockman, S. 1913 Report of the departmental committee appointed by the board of agriculture and fisheries to inquire into epizootic abortion III. Abortion in Sheep. HMSO. London. UK. Stockman, S.

McGuigan, F. E. and Ralston, S. H. 2002, Single nucleotide polymorphism detection: allelic discrimination using TaqMan, Psychiatr. Genet., vol. 12, no. 3, 133-136.

McKillip, J. L., Jaykus, L. A., and Drake, M. 1999, Nucleic acid persistence in heat-killed *Escherichia coli* O157:H7 from contaminated skim milk, J. Food. Prot., vol. 62, no. 8, 839-844.

Mclauchlin, J., Ripabelli, G., Brett, M. M., and Threlfall, E. J. 2000, Amplified fragment length polymorphism (AFLP) analysis of *Clostridium perfringens* for epidemiological typing, Int. J. Food Microbiol., vol. 56, no. 1, 21-28.

McOrist, A. L., Jackson, M., and Bird, A. R. 2002, A comparison of five methods for extraction of bacterial DNA from human faecal samples, J. Microbiol. Methods., vol. 50, 131-139.

McQuiston, J. R., Parrenas, R., Ortiz-Rivera, M., Gheesling, L., Brenner, F., and Fields, P. I. 2004, Sequencing and comparative analysis of flagellin genes *fliC*, *fljB*, and *flpA* from *Salmonella*, J. Clin. Microbiol., vol. 42, no. 5, 1923-1932.

McSweegan, E. and Walker, R. I. 1986, Identification and characterization of two *Campylobacter jejuni* adhesins for cellular and mucous substrates, Infect. Immun., vol. 53, no. 1, 141-148.

Megraud, F., Boudraa, G., Bessaoud, K., Bensid, S., Dabis, F., Soltana, R., and Touhami, M. 1990, Incidence of campylobacter infection in infants in western Algeria and the possible protective role of breast feeding, Epidemiol. Infect., vol. 105, no. 1, 73-78.

Megraud, F. and Serceau, R. 1990, Search for *Campylobacter* species in the public water supply of a large urban community, Zentralbl.Hyg.Umweltmed., vol. 189, no. 6, 536-542.

Meinersmann, R. J., Helsel, L. O., Fields, P. I., and Hiett, K. L. 1997, Discrimination of *Campylobacter jejuni* isolates by *fla* gene sequencing, J. Clin. Microbiol., vol. 35, no. 11, 2810-2814.

Metherell, L. A., Logan, J. M., and Stanley, J. 1999, PCR-enzyme-linked immunosorbent assay for detection and identification of *Campylobacter* species: application to isolates and stool samples, J. Clin. Microbiol., vol. 37, no. 2, 433-435.

Miesfield, R. L. 1999, "Characterisation of Genomic DNA," in *Applied Molecular Genetics*, John Wiley and sons inc, pp. 83-114.

Misawa, N., Hirayama, K., Itoh, K., and Takahashi, E. 1995, Detection of alpha- and beta-hemolytic-like activity from *Campylobacter jejuni*, J. Clin. Microbiol., vol. 33, no. 3, 729-731.

Misawa, N., Kawashima, K., Kawamoto, H., and Kondo, F. 2002, Development of a combined filtration-enrichment culture followed by a one-step duplex PCR technique for the rapid detection of *Campylobacter jejuni* and *C. coli* in human faecal samples, J.Med.Microbiol., vol. 51, no. 1, 86-89.

Mohran, Z. S., Arthur, R. R., Oyofo, B. A., Peruski, L. F., Wasfy, M. O., Ismail, T. F., and Murphy, J. R. 1998, Differentiation of campylobacter isolates on the basis of sensitivity to boiling in water as measured by PCR-detectable DNA, Appl.Environ.Microbiol., vol. 64, 363-365.

Monstein, H., Nikpour-Badr, S., and Jonasson, J. 2001, Rapid molecular identification and subtyping of *Helicobacter pylori* by pyrosequencing of the 16S rDNA variable V1 and V3 regions, FEMS Microbiol.Lett., vol. 199, no. 1, 103-107.

Monteville, M. R. and Konkel, M. E. 2002, Fibronectin-facilitated invasion of T84 eukaryotic cells by *Campylobacter jejuni* occurs preferentially at the basolateral cell surface, Infect.Immun., vol. 70, no. 12, 6665-6671.

Monteville, M. R., Yoon, J. E., and Konkel, M. E. 2003, Maximal adherence and invasion of INT 407 cells by *Campylobacter jejuni* requires the CadF outer-membrane protein and microfilament reorganization, Microbiology, vol. 149, no. Pt 1, 153-165.

Moore, C. M., Sheldon, B. W., and Jaykus, L. A. 2003, Transfer of *Salmonella* and *Campylobacter* from stainless steel to romaine lettuce, J.Food Prot., vol. 66, no. 12, 2231-2236.

Moore, J., Caldwell, P., and Millar, B. 2001, Molecular detection of *Campylobacter* spp. in drinking, recreational and environmental water supplies, Int.J.Hyg.Environ.Health, vol. 204, no. 2-3, 185-189.

Moore, J. E. and Murphy, P. G. 2003, Comparison of human and porcine strains of *Campylobacter coli*, Ir.J Med Sci, vol. 172, no. 2, 89-90.

Moore, M. A., Blaser, M. J., Perez-Perez, G. I., and O'Brien, A. D. 1988, Production of a Shiga-like cytotoxin by *Campylobacter*, Microb.Pathog., vol. 4, no. 6, 455-462.

Moorhead, S. M., Dykes, G. A., and Cursons, R. T. 2003, An SNP-based PCR assay to differentiate between *Listeria monocytogenes* lineages derived from phylogenetic analysis of the *sig*B gene, J Microbiol.Methods, vol. 55, no. 2, 425-432.

Moran, A. P. and Penner, J. L. 1999, Serotyping of *Campylobacter jejuni* based on heat-stable antigens:Relevance, molecular basis and implications in pathogenesis, J. Appl. Microbiol., vol. 86, 361-377.

Morin, P. A., Saiz, R., and Monjazeb, A. 1999, High-throughput single nucleotide polymorphism genotyping by fluorescent 5' exonuclease assay, Biotechniques, vol. 27, no. 3, 538-40, 542, 544.

Morooka, T., Umeda, A., Fujita, M., Matano, H., Fujimoto, S., Yukitake, K., Amako, K., and Oda, T. 1996, Epidemiologic application of pulsed-field gel electrophoresis to an outbreak of *Campylobacter fetus* meningitis in a neonatal intensive care unit, Scand. J. Infect.Dis., vol. 28, no. 3, 269-270.

Murphy, P. G. 1993, "Multilocus Enzyme Electrophoresis," in *New Techniques in Food and Beverage Microbiology*, Society for Applied Bacteriology, pp. 211-221.

Nachamkin, I., Bohachick, K., and Patton, C. M. 1993, Flagellin gene typing of *Campylobacter jejuni* by restriction fragment length polymorphism analysis, J. Clin. Microbiol., vol. 31, no. 6, 1531-1536.

Nachamkin, I., Engberg, J., Gutacker, M., Meinersman, R. J., Li, C. Y., Arzate, P., Teeple, E., Fussing, V., Ho, T. W., Asbury, A. K., Griffin, J. W., McKhann, G. M., and Piffaretti, J. C. 15-7-2001, Molecular population genetic analysis of *Campylobacter jejuni* HS:19 associated with Guillain-Barre syndrome and gastroenteritis, J. Infect. Dis., vol. 184, no. 2, 221-226.

Nair, S., Schreiber, E., Thong, K-L., Pang, T., and Altwegg, W. 2000, Genotypic characterisation of *Salmonella typhi* by amplified fragment length polymorphism fingerprinting provides increased discrimination as compared to pulsed field gel electrophoresis and ribotyping, J. Microbiol. Methods, vol. 41, 35-43.

Nawaz, M. S., Khan, S. A., Khan, A. A., Nayak, R., Steele, R., Paine, D., and Jones, R. 2003a, Molecular characterization of fluoroquinolone-resistant *Campylobacter* spp. isolated from poultry, Poultry Sci., vol. 82, no. 2, 251-258.

Nawaz, M. S., Wang, R. F., Khan, S. A., and Khan, A. A. 2003b, Detection of *gal*E gene by polymerase chain reaction in campylobacters associated with Guillain-Barre syndrome, Mol. Cell Probes, vol. 17, no. 6, 313-317.

Newell, D. 2002, The ecology of *Campylobacter jejuni* in avian and human hosts and in the environment, Int. J. Infect. Dis., vol. 6, 3S16-3S21.

Newell, D. G. 2001, Animal models of *Campylobacter jejuni* colonization and disease and the lessons to be learned from similar *Helicobacter pylori* models, Symp. Ser. Soc. Appl. Microbiol. no. 30, 57S-67S.

Newell, D. G. and Fearnley, C. 2003, Sources of Campylobacter colonization in broiler chickens, Appl. Environ. Microbiol., vol. 69, no. 8, 4343-4351.

Newell, D. G., Frost, J. A., Duim, B., Wagenaar, J. A., Madden, R. H., Van der Plas, J., & On, S. L. W. 2000, "New developments in the subtyping of *Campylobacter* species," in *Campylobacter*, 2 edn, I. B. M. J. Nachampkin, ed., American Society for Microbiology., Washington, pp. 27-44.

Newell, D. G., Shreeve, J. E., Toszeghy, M., Domingue, G., Bull, S., Humphrey, T., and Mead, G. 2001, Changes in the carriage of *Campylobacter* strains by poultry carcasses during processing in abattoirs, Appl. Environ. Microbiol., vol. 67, no. 6, 2636-2640.

Ng, L. K., Kingombe, C. I., Yan, W., Taylor, D. E., Hiratsuka, K., Malik, N., and Garcia, M. M. 1997, Specific detection and confirmation of *Campylobacter jejuni* by DNA hybridization and PCR, Appl. Environ. Microbiol., vol. 63, no. 11, 4558-4563.

Nielsen, E. M. 2002, Occurrence and strain diversity of thermophilic campylobacters in cattle of different age groups in dairy herds, Lett. Appl. Microbiol., vol. 35, no. 1, 85-89.

Nogva, H. K., Bergh, A., Holck, A., and Rudi, K. 2000a, Application of the 5' Nuclease PCR assay in evaluation and development of methods for quantification and detection of *Campylobacter jejuni*, Appl. Environ. Microbiol., vol. 66, no. 9, 4029-4036.

Nogva, H. K., Dromtorp, S. M., Nissen, H., and Rudi, K. 2003, Ethidium monoazide for DNA-based differentiation of viable and dead bacteria by 5'-nuclease PCR, Biotechniques, vol. 34, no. 4, 804-3.

Nogva, H. K., Rudi, K., Naterstad, K., Holck, A., and Lillhaug, D. 2000b, Application of 5'-Nuclease PCR for quantitative detection of *Listeria monocytogenes* in pure cultures, water, skim milk and unpasteurised whole milk., Appl. Environ. Microbiology, vol. 66, no. 10, 4266-4271.

Nylen, G., Dunstan, F., Palmer, S. R., Andersson, Y., Bager, F., Cowden, J., Feierl, G., Galloway, Y., Kapperud, G., Megraud, F., Molbak, K., Petersen, L. R., and Ruutu, P. 2002, The seasonal distribution of campylobacter infection in nine European countries and New Zealand, Epidemiol. Infect., vol. 128, no. 3, 383-390.

O'Sullivan, N., Benjamin, J., and Skirrow, M. B. 1990, Lectin typing of campylobacter isolates, J. Clin. Pathol., vol. 43, no. 11, 957-960.

Obiri-Danso, K. and Jones, K. 1999, Distribution and seasonality of microbial indicators and thermophilic campylobacters in two freshwater bathing sites on the river Lune in northwest England, J. Appl. Microbiol., vol. 87, no. 6, 822-832.

Oelschlaeger, T. A., Guerry, P., and Kopecko, D. J. 1993, Unusual microtubule-dependent endocytosis mechanisms triggered by *Campylobacter jejuni* and *Citrobacter freundii*, Proc. Natl. Acad. Sci. U.S.A, vol. 90, no. 14, 6884-6888.

Olsen, S. J., Hansen, G. R., Barlett, L., Fitzgerald, C., Sonder, A., Manjrekar, R., Riggs, T., Kim, J., Flahart, R., Pezzino, G., and Swerdlow, D. L. 2001, An outbreak of *Campylobacter jejuni* infections associated with food handler contamination: The use of pulsed field gel electrophoresis, J. Infect. Dis., vol. 183, 164-167.

On, S. L W and Jordan, P. J. 2003, Evaluation of 11 PCR assays for species-level identifiaction of *Campylobacter jejuni* and *Campylobacter coli*, J. Clin. Microbiol., vol. 41, no. 1, 330-336.

On, S. L. 1996a, Identification methods for campylobacters, helicobacters, and related organisms. Clin. Microbiol. Rev., vol. 9, no. 3, 405-422.

On, S. L. and Harrington, C. S. 1-12-2000, Identification of taxonomic and epidemiological relationships among *Campylobacter* species by numerical analysis of AFLP profiles, FEMS Microbiol. Lett., vol. 193, no. 1, 161-169.

On, S. L. and Harrington, C. S. 2001, Evaluation of numerical analysis of PFGE-DNA profiles for differentiating *Campylobacter fetus* subspecies by comparison with phenotypic, PCR and 16S rDNA sequencing methods, J. Appl. Microbiol., vol. 90, no. 2, 285-293.

On, S. L. W. 1996b, Identification Methods for *Campylobacters*, *Helicobacters* and Related Organisms, Clin. Microbiol. Rev., vol. 9, no. 3, 405-422.

Ono, K., Kurazono, T., Niwa, H., and Itoh, K. 2003, Comparison of three methods for epidemiological typing of *Campylobacter jejuni* and *C. coli*, Curr. Microbiol., vol. 47, no. 5, 364-371.

Owen, R. J., Costas, M., Sloss, L., and Bolton, F. J. 1988, Numerical analysis of electrophoretic protein patterns of *Campylobacter laridis* and allied thermophilic campylobacters from the natural environment, J. Appl. Bacteriol., vol. 65, no. 1, 69-78.

Oyarzabal, O. A., Wesley, I. V., Barbaree, J. M., Lauerman, L. H., and Conner, D. E. 1997a, Specific detection of *Campylobacter lari* by PCR, J. Microbiol. Methods, vol. 29, 97-102.

Oyarzabal, O. A., Wesley, I. V., Harmon, K. M., Schroeder-Tucker, L., Barbaree, J. M., Lauerman, L. H., Backert, S., and Conner, D. E. 1997b, Specific identification of *Campylobacter fetus* by PCR targeting variable regions of the 16S rDNA, Vet. Microbiol., vol. 58, no. 1, 61-71.

Panet, A. and Khorana, G. 1974, Studies on polynucleotides. The linkage of deoxyribonucleotide templates to cellulose and its use in their replication, J. Biol. Chem., vol. 249, no. 16, 5213-5221.

Park, S. F. 2002, The physiology of *Campylobacter* species and its relevance to their role as foodborne pathogens, Int. J. Food Microbiol., vol. 74, 177-188.

Park, S. F., Purdy, D., and Leach, S. 2000, Localized reversible frameshift mutation in the *flh*A gene confers phase variability to flagellin gene expression in *Campylobacter coli*, J. Bacteriol., vol. 182, no. 1, 207-210.

Parkhill, J., Wren, B. W., Mungall, K., Ketley, J.M., Churher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, S., Jagels, K., Karylshev, A. V., Moule, S., Pallen, M. J., Penn, C. W., Quail, M. A., Rajendream, M. A., Van Vilet, A. H. M., Whithead, S., and Barrell, B. G. 2000, The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences, Nature, vol. 403, 665-668.

Pearson, A. D., Greenwood, M., Healing, T. D., Rollins, D., Shahamat, M., Donaldson, J., and Colwell, R. R. 1993, Colonization of broiler chickens by waterborne *Campylobacter jejuni*, Appl. Environ. Microbiol., vol. 59, no. 4, 987-996.

Pedersen, K. and Wedderkopp, A. 2003, Resistance to quinolones in *Campylobacter jejuni* and *Campylobacter coli* from Danish broilers at farm level, J. Appl. Microbiol., vol. 94, no. 1, 111-119.

Pei, Z. and Blaser, M. J. 1993, PEB1, the major cell-binding factor of *Campylobacter jejuni*, is a homolog of the binding component in gram-negative nutrient transport systems, J. Biol. Chem., vol. 268, no. 25, 18717-18725.

Pei, Z., Burucoa, C., Grignon, B., Baqar, S., Huang, X. Z., Kopecko, D. J., Bourgeois, A. L., Fauchere, J. L., and Blaser, M. J. 1998, Mutation in the *peb*1A locus of *Campylobacter jejuni* reduces interactions with epithelial cells and intestinal colonization of mice, Infect. Immun., vol. 66, no. 3, 938-943.

Penner, J. M. and Hennessey, J. N. 1980, Passive haemaglutination technique for serotyping *Campylobacter fetus* subspecies *jejuni* on the basis of soluble heat stable antigens., J.Clin. Microbiol., vol. 12, 732-737.

Perko-Makela, P., Hakkinen, M., Honkanen-Buzalski, T., and Hanninen, M. L. 2002, Prevalence of campylobacters in chicken flocks during the summer of 1999 in Finland, Epidemiol. Infect., vol. 129, no. 1, 187-192.

Perutz, M. F. 1983, Species adaptation in a protein molecule, Mol. Biol. Evol., vol. 1, no. 1, 1-28.

Peters, T. M., Maguire, C., Threlfall, E. J., Fisher, I. S., Gill, N., and Gatto, A. J. 2003, The Salmgene project - a European collaboration for DNA fingerprinting for, Euro. Surveill., vol. 8, no. 2, 46-50.

Peterson, M. C. 2003, *Campylobacter jejuni* enteritis associated with consumption of raw milk, J. Environ. Health, vol. 65, no. 9, 20-1, 24, 26.

Pickett, C. L. 2000, "Campylobacter toxins and their role in pathogenesis," in *Campylobacter*, I. Nachamkin & J. M. Blaser, eds., American Society for Microbiology, Washington D.C, pp. 179-190.

Pickett, C. L., Auffenberg, T., Pesci, E. C., Sheen, V. L., and Jusuf, S. S. 1992, Iron acquisition and hemolysin production by *Campylobacter jejuni*, Infect. Immun., vol. 60, no. 9, 3872-3877.

Pickett, C. L., Pesci, E. C., Cottle, D. L., Russell, G., Erdem, A. N., and Zeytin, H. 1996, Prevalence of cytolethal distending toxin production in *Campylobacter jejuni* and relatedness of *Campylobacter* sp. *cdt*B gene, Infect.Immun., vol. 64, no. 6, 2070-2078.

Plastridge, W. N., Williams, L. F., and Pertrie, D. 1947, Vibrionic Abortion in Cattle, Am. J. Vet. Res., vol. 8, 178-183.

Popoff, M. Y., Bockemuhl, J., Brenner, F. W., and Gheesling, L. L. 2001, Supplement 2000 (no. 44) to the Kauffmann-White scheme, Res. Microbiol., vol. 152, no. 10, 907-909.

Porter, I. A., and Reid, T. M. 1980, A milk-borne outbreak of *Campylobacter* infection, J. Hyg. (Lon). vol. 84, no. 3, 415-419.

Potter, R. C., Kaneene, J. B., and Hall, W. N. 2003, Risk factors for sporadic *Campylobacter jejuni* infections in rural michigan: a prospective case-control study, Am. J. Public Health, vol. 93, no. 12, 2118-2123.

Ramaswamy, S. V., Reich, R., Dou, S. J., Jasperse, L., Pan, X., Wanger, A., Quitugua, T., and Graviss, E. A. 2003, Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*, Antimicrob. Agents Chemother., vol. 47, no. 4, 1241-1250.

Ranade, K., Chang, M. S., Ting, C. T., Pei, D., Hsiao, C. F., Olivier, M., Pesich, R., Hebert, J., Chen, Y. D., Dzau, V. J., Curb, D., Olshen, R., Risch, N., Cox, D. R., and Botstein, D. 2001, High-throughput genotyping with single nucleotide polymorphisms, Genome Res., vol. 11, no. 7, 1262-1268.

Raupach, J. C. and Hundy, R. L. 2003, An outbreak of *Campylobacter jejuni* infection among conference delegates, Commun. Dis. Intell., vol. 27, no. 3, 380-383.

Read, T. D., Salzberg, S. L., Pop, M., Shumway, M., Umayam, L., Jiang, L., Holtzapple, E., Busch, J. D., Smith, K. L., Schupp, J. M., Solomon, D., Keim, P., and Fraser, C. M. 2002, Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*, Science, vol. 296, no. 5575, 2028-2033.

Refregier-Petton, J., Rose, N., Denis, M., and Salvat, G. 2001, Risk factors for *Campylobacter* spp. contamination in french broiler-chicken flocks at the end of the rearing period, Prev. Vet. Med., vol. 19;50, no. 1-2, 89-100.

Reierson, J., Briem, H., Harddardottir, H., Gunnarson, E., Georgsson, F., & Kristinsson, K. G. 2001, "Human campylobacteriosis epidemic in Iceland 1998-2000 and the effect of interventions aimed at Poultry and Animals", Int. J. Med. Microbiol. vol. 291, suppl 31, 153.

Reinders, R. D., Barna, A., Lipman, L. J., and Bijker, P. G. 2002, Comparison of the sensitivity of manual and automated immunomagnetic separation methods for detection of Shiga toxin-producing *Escherichia coli* O157:H7 in milk, J. Appl. Microbiol., vol. 92, no. 6, 1015-1020.

Reiser, A., Geyer, M, Van Miltenberg, R., Nauck, M., and Tabiti, K. 1999, Mutation detection using multi-color detection on the Lightcycler system, Biochemica, vol. 2, 12-15.

Ribot, E. M., Fitzgerald, C., Kubota, K., Swaiminathan, B., and Barrett, T. 2001, Rapid pulsed field gel electrophoresis protocol for subtyping of *Campylobacter jejuni*, J. Clin. Microbiol., vol. 39, no. 5, 1889-1894.

Richardson and Park, S. F. 1995, Enterochelin acquisition in *Campylobacter coli*: Characterisation of components of a binding-protein-dependent transport system, Microbiology, vol. 141, 3181-3191.

Richardson, J. F., Frost, J. A., Kramer, J. M., Thwaites, R. T., Bolton, F. E., Wareing, D. R., and Gordon, J. A. 2002, Coinfection with *Campylobacter* species: and epidemiological problem, J. Appl. Microbiol., vol. 91, 206-211.

Riordan, T., Humphrey, T. J., and Fowles, A. 1993, A point source outbreak of campylobacter infection related to bird-pecked milk, Epidemiol. Infect., vol. 110, no. 2, 261-265.

Ripabelli, G., McLauchin, J., and Threlfall, E. J. 2000a, Amplified fragment length polymorphism (AFLP) analysis of *Listeria monocytogenes*, Syst. Appl. Microbiol., vol. 23, no. 1, 132-136.

Ripabelli, G., Mclauchlin, J., Mithani, V., and Threlfall, E. J. 2000b, Epidemiological typing of *Bacillus cereus* by amplified fragment length polymorphism, Lett. Appl. Microbiol., vol. 30, no. 5, 358-363.

Rivoal, K., Denis, M., Salvat, G., Colin, P., and Ermel, G. 1999, Molecular characterization of the diversity of *Campylobacter* spp. isolates collected from a poultry slaughterhouse: analysis of cross-contamination, Lett. Appl. Microbiol., vol. 29, no. 6, 370-374.

Robertson, G. A., Thiruvenkataswamy, V., Shilling, H., Price, E. P., Huygens, F., Henskens, F. A., and Giffard, P. M. 2004, Identification and interrogation of highly informative single nucleotide polymorphism sets defined by bacterial multilocus sequence typing databases, J. Med. Microbiol., vol. 53, no. 1, 35-45.

Robinson, D. A. and Jones, D. M. 1981, Milk-borne campylobacter infection, Br. Med. J. (Clin Res.Ed), vol. 282, no. 6273, 1374-1376.

Rock, J. D., Van Vilet, A. H. M., and Ketley, J. M. 2001, Haemin uptake in *Campylobacter jejuni*, Int. J. Med. Microbiol., vol. 291, no. Suppl 31, 125.

Romling, U. and Tummler, B. 2000, Achieving 100% typeability of *Pseudomonas aeruginosa* by pulsed-field gel electrophoresis, J. Clin. Microbiol., vol. 38, no. 1, 464-465.

Roop, R. M., Smibert, R. M., Johnson, J. L., and Krieg, N. R. 1984, Differential characteristics of catalase-positive campylobacters correlated with DNA homology groups, Can. J. Microbiol., vol. 30, no. 7, 938-951.

Rudi, K., Nogva, H. K., Moen, B., Nissen, H., Bredholt, S., Moretro, T., Naterstad, K., and Holck, A. 2002, Development and application of new nucleic acid-based technologies for microbial community analyses in foods, Int. J. Food Microbiol., vol. 78, no. 1-2, 171-180.

Rudi, K. and Holck, A. L. 2003, Real-time closed tube single nucleotide polymorphism (SNP) quantification in pooled samples by quencher extension (QEXT), Nucleic Acids Res., vol. 31, no. 19, 117.

Russell, P. J. 2003, "Genome Analysis," in *Essential Genetics*, Benjamin Cummings, San Fransico, p. 30.

Sahin, O., Kobalka, P., and Zhang, Q. 2003, Detection and survival of Campylobacter in chicken eggs, J. Appl. Microbiol., vol. 95, no. 5, 1070-1079.

Said, B., Wright, F., Nichols, G. L., Reacher, M., and Rutter, M. 2003, Outbreaks of infectious disease associated with private drinking water supplies in England and Wales 1970-2000, Epidemiol. Infect., vol. 130, no. 3, 469-479.

Sails, A. D. 2002. Detection of *Campylobacter jejuni* and *Campylobacter coli* in high risk foods and environmental waters using the polymerase chain reaction. PhD thesis, University of Central Lancashire. pp. 63.

Sails, A. D., Bolton, F. J., Fox, A. J., Wareing, D. R., and Greenway, D. L. 1998, A reverse transcriptase polymerase chain reaction assay for the detection of thermophilic *Campylobacter spp*, Mol. Cell Probes, vol. 12, no. 5, 317-322.

Sails, A. D., Bolton, F. J., Fox, A. J., Wareing, D. R. A., and Greenway, D. L. A. 2002, Detection of *Campylobacter jejuni* and *Campylobacter coli* in environmental waters by PCR enzyme-linked immunosorbent assay, Appl. Environ. Microbiol., vol. 68, 1319-1324.

Sails, A. D., Fox, A. J., Bolton, F. E., Wareing, D., Greenway, D. L., and Borrow, R. 2001, Development of a PCR ELISA assay for the identification of *Campylobacter jejuni* and *Campylobacter coli*, Mol. Cell Probes, vol. 15, no. 5, 291-300.

Sails, A. D., Fox, A. J., Bolton, F. J., Wareing, D. R. A., and Greenway, D. L. A. 2003a, A real-time PCR assay for the detection of *Campylobacter jejuni* in foods after enrichment culture, Appl. Environ. Microbiol., vol. 69, 1383-1390.

Sails, A. D., Swaminathan, B., and Fields, P. I. 2003b, Clonal complexes of *Campylobacter jejuni* identified by multilocus sequence typing correlate with strain associations identified by multilocus Enzyme electrophoresis, J. Clin. Microbiol., vol. 41, no. 9, 4058-4067.

Sails, A. D., Swaminathan, B., and Fields, P. I. 2003c, Utility of multilocus sequence typing as an epidemiological tool for investigation of outbreaks of gastroenteritis caused by *Campylobacter jejuni*, J. Clin. Microbiol., vol. 41, no. 10, 4733-4739.

Salama, S. M., Bolton, F. J., and Hutchinson, D. N. 1990, Application of a new phage typing sheme to campylobacters isolated during outbreaks, Epidemiol. Infect., vol. 104, 405-411.

Salcedo, C., Arreaza, L., Alcala, B., de la, Fuente L., and Vazquez, J. A. 2003, Development of a multilocus sequence typing method for analysis of *Listeria monocytogenes* clones, J. Clin. Microbiol., vol. 41, no. 2, 757-762.

Salyers, A. A. & Witt, D. D. 1994, "Virulence Factors that Promote Colonisation," in *Molecular Pathogenesis*, 1 edn, American Society for Microbiology, N.Y. Washington D.C., pp. 30-46.

Santamaria, J. and Toranzos, G. A. 2003, Enteric pathogens and soil: a short review, Int. Microbiol., vol. 6, no. 1, 5-9.

Sasaki, Y., Fujisawa, T., Ogikubo, K., Ohzono, T., Ishihara, K., and Takahashi, T. 2003, Characterization of *Campylobacter lanienae* from pig feces, J. Vet. Med. Sci., vol. 65, no. 1, 129-131.

Saunders, N. A., Harrison, T. G., Haththotuwa, A., Kachwalla, N., and Taylor, A. G. 1990, A method for typing strains of *Legionella pneumophila* serogroup 1 by analysis of restriction fragment length polymorphisms, J. Med. Microbiol., vol. 31, no. 1, 45-55.

Schouls, L. M., Reulen, S., Duim, B., Wagenaar, J. A., Willems, R. J., Dingle, K. E., Colles, F. M., and Van Embden, J. D. 2003, Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination, J. Clin. Microbiol., vol. 41, no. 1, 15-26.

Sebald, M. and Vernon, M. 1963, DNA base content in the classification of Vibrios, Ann.Inst.pasteur, vol. 105, 897-910.

Sheridan, G. E., Masters, C. I., Shallcross, J. A., and MacKey, B. M. 1998, Detection of mRNA by reverse transcription-PCR as an indicator of viability in *Escherichia coli* cells, Appl. Environ. Microbiol., vol. 64, no. 4, 1313-1318.

Shi, F., Chen, Y. Y., Wassenaar, T. M., Woods, W. H., Coloe, P. J., and Fry, B. N. 2002, Development and application of a new scheme for typing *Campylobacter jejuni* and *Campylobacter coli* by PCR-based restriction fragment length polymorphism analysis, J. Clin. Microbiol., vol. 40, no. 5, 1791-1797.

Simmon, K. E., Steadman, D. D., Durkin, S., Baldwin, A., Jeffrey, W. H., Sheridan, P., Horton, R., and Shields, M. S. 2004, Autoclave method for rapid preparation of bacterial PCR-template DNA, J. Microbiol. Methods, vol. 56, no. 2, 143-149.

Skirrow, M. B. 1977, Campylobacter enteritis: a "new" disease, Br. Med. J., vol. 2, no. 6078, 9-11.

Skirrow, M. B. & Blaser, M. J. 2000, "Clinical Aspects of *Campylobacter* Infection," in *Campylobacter*, 2 edn, I. B. M. J. Nachampkin, ed., American Society for microbiology, Washington D.C., pp. 69-88.

Skirrow, M. B., Jones, D. M., Sutcliffe, E., and Benjamin, J. 1993, *Campylobacter* bacteraemia in England and Wales, 1981-91, Epidemiol. Infect., vol. 110, no. 3, 567-573.

Smith, S. I., Olukoya, D. K., Fox, A. J., and Coker, A. O. 2000a, Deoxyribonucleic acid restriction patterns in *Campylobacter* species: a comparison with Penner serotype., Br. J. Biomed. Sci.,vol. 57, 137-141.

Smith, S. I., Olukoya, D. K., Fox, A. J., and Coker, A. O. 2000b, Genotyping of clinical and chicken isolates of *Campylobacter jejuni* and *Campylobacter coli*, Cytobios, vol. 103, 91-101.

Sokurenko, E. V., Chesnokova, V., Dykhuizen, D. E., Ofek, I., Wu, X. R., Krogfelt, K. A., Struve, C., Schembri, M. A., and Hasty, D. L. 1998, Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin, Proc. Natl. Acad. Sci., U.S.A, vol. 95, no. 15, 8922-8926.

Spratt, B. 2002, Bacterial Populations Structures and Molecular typing - an overview. International Workshop on Multilocus Sequence Typing, Hinxton. U.K. 1-3/12/2002.

Stanley, K. and Jones, K. 2003, Cattle and sheep farms as reservoirs of Campylobacter, J. Appl. Microbiol., vol. 94 Suppl:104S-113S.

Stanley, K. N., Wallace, J. S., Currie, J. E., Diggle, P. J., and Jones, K. 1998a, Seasonal Variation of thermophillic campylobacters in lambs at slaughter, J. Appl. Microbiol., vol. 84, 405-411.

Stanley, K. N., Wallace, J. S., Currie, J. E., Diggle, P. J., and Jones, K. 1998b, The seasonal variation of thermophilic campylobacters in beef cattle, dairy cattle and calves, J. Appl. Microbiol., vol. 85, no. 3, 472-480.

Stapleton, K., Iqbal, J., Toszeghy, M., Newell, D., and Ridley, A. 2003, Development of a rapid PCR based Lightcycler assay to speciate *Campylobacter jejuni* and *C.coli* isolates from veterinary faecal and caecal samples, Int. J. Med. Microbiol., vol. 293, Supplement 35, no. 4.

Stermann, M., Bohrssen, A., Diephaus, C., Maass, S., and Bange, F. C. 2003, Polymorphic nucleotide within the promoter of nitrate reductase (NarGHJI) is specific for *Mycobacterium tuberculosis*, J. Clin. Microbiol., vol. 41, no. 7, 3252-3259.

Stonnet, V. and Guesdon, J-L. 1993, *Campylobacter jejuni*: specific oligonucleotides and DNA probes for use in Polymerase Chain Reaction based diagnosis, FEMS Immunol. Med. Microbiol., vol. 7, 337-344.

Stuart, J., Sufi, F., McNulty, C., and Park, P. 1997, Outbreak of campylobacter enteritis in a residential school associated with bird pecked bottle tops, Commun. Dis. Rep. CDR Rev., vol. 7, no. 3, R38-R40.

Stucki, U., Frey, J., Nicolet, J., and Burnens, A. P. 1995, Identification of *Campylobacter jejuni* on the basis of a species-specific gene that encodes a membrane specific protein, J. Clin. Microbiol., vol. 33, no. 4, 855-859.

Suerbaum, S., Lohrengel, M., Sonnevend, A., Ruberg, F., and Kist, M. 2001, Allelic diversity and recombination in *Campylobacter jejuni*, J. Bacteriol., vol. 183, no. 8, 2553-2559.

Talibart, R., Denis, M., Castillo, A., Cappelier, J. M., and Ermel, G. 2000, Survival and recovery of viable but noncultivable forms of *Campylobacter* in aqueous microcosm, Int. J. Food Microbiol., vol. 55, no. 1-3, 263-267.

Tam, C. C., Rodrigues, L. C., and O'Brien, S. J. 2003, Guillain-Barre syndrome associated with *Campylobacter jejuni* infection in England 2000-2001, Clin. Infect. Dis., vol. 37, 307-310.

Taylor, D. E., Eaton, M., Yan, W., and Chang, N. 1992, Genome maps of *Campylobacter jejuni* and *Campylobacter coli*, J. Bacteriol., vol. 174, no. 7, 2332-2337.

Taylor, D. N. 1992, "Campylobacter infections in developing countries," in *Campylobacter jejuni:Current status and future needs*, I. Nachamkin, M. J. Blaser, & L. S. Tompkins, eds., American Society for Microbiology, pp. 20-30.

Taylor, J. W. and Fisher, M. C. 2003, Fungal Multilocus sequence typing- it's not just for bacteria, Curr. Opin. Microbiol., vol. 6, 351-356.

Tenkate, T. D. and Stafford, R. J. 2001, Risk factors for campylobacter infection in infants and young children: a matched case-control study, Epidemiol. Infect., vol. 127, no. 3, 399-404.

Tholozan, J. L., Cappelier, J. M., Tissier, J. P., Delattre, G., and Federighi, M. 1999, Physiological characterization of viable but non-culturable *Campylobacter jejuni* cells, Appl. Environ. Microbiol., vol. 65, no. 3, 1110-1116.

Thunberg, R. L., Tran, T. T., and Walderhaug, M. O. 2000, Detection of thermophilic *Campylobacter* spp. in blood-free enriched samples of inoculated foods by the polymerase chain reaction, J. Food Prot., vol. 63, no. 3, 299-303.

Thwaites, R. T. and Frost, J. A. 1997, Drug resistance in *Campylobacter jejuni, C.coli*, and *C.lari* isolated from humans in North West England and Wales 1997, J. Clin. Path., vol. 52, 812-814.

Totten, P. A., Patton, F. C., Tenover, F. C., Barrett, T. J., Stamm, W. E., Steigerwalt, A. G., Lin, J. Y., Holmes, K. K., and Brenner, D. J. 1987, Prevalence and characterisation of Hippurate-negative *Campylobacter jejuni* in King Country, Washington, J. Clin. Microbiol., vol. 25, 1747-1752.

Turabelidze, D., Kotetishvili, M., Kreger, A., Morris, J. G., Jr., and Sulakvelidze, A. 2000, Improved pulsed-field gel electrophoresis for typing vancomycin-resistant enterococci, J. Clin. Microbiol., vol. 38, no. 11, 4242-4245.

Tzanakaki, G., Urwin, R., Musilek, M., Kriz, P., Kremastinou, J., Pangalis, A., Blackwell, C. C., and Maiden, M. C. 2001, Phenotypic and genotypic approaches to characterization of isolates of *Neisseria meningitidis* from patients and their close family contacts, J. Clin. Microbiol., vol. 39, no. 4, 1235-1240.

Unnerstad, H., Ericsson, H., Alderborn, A., Tham, W., Danielsson-Tham, M. L., and Mattsson, J. G. 2001, Pyrosequencing as a method for grouping of *Listeria monocytogenes* strains on the basis of single-nucleotide polymorphisms in the *inl*B gene, Appl. Environ. Microbiol., vol. 67, no. 11, 5339-5342.

Urwin, R. and Maiden, M. C. 2003, Multi-locus sequence typing: a tool for global epidemiology, Trends Microbiol., vol. 11, no. 10, 479-487.

Van Camp, G., Fierens, H., Van Damme, P., Goossens, H., Huyghebaert, A., and De Watcher, R. 1993, Identification of enteropathogenic *Campylobacter* species by oligonucleotide probes and polymerase chain reaction based on 16S rRNA genes, Sys. Appl. Microbiol., vol. 16, 30-36.

Van Doorn, L-J, Gisendorf, B. A. J., Bax, R., Van der Zeijst, B. A. M., Vandamme, P., and Quint, W. G. V. 1997, Molecular Discrimination between *Campylobacter jejuni, Campylobacter coli, campylobacter lari* and *campylobacter upsaliensis* by polymerase chain reaction based on a novel putative GTPase gene, Mol. Cell. Probes, vol. 11, 177-185.

van Leeuwen, W. B., Jay, C., Snijders, S., Durin, N., Lacroix, B., Verbrugh, H. A., Enright, M. C., Troesch, A., and van Belkum, A 2003, Multilocus Sequence Typing of *Staphylococcus aureus* with DNA array technology, J. Clin. Microbiol., vol. 41, no. 7, 3323-3326.

van Vliet, A. H., Ketley, J. M., Park, S. F., and Penn, C. W. 2002, The role of iron in campylobacter gene regulation, metabolism and oxidative stress defense., FEMS Microbiol. Rev., vol. 26, no. 2, 173-186.

Vandamme, P, Van Doorn, L. J., Al Rashid, S. T., Quint, W. G. V., Van der Plas, J., Chan, V. L., and On, S. L. W. 1997, *Campylobacter hyoilei* Alderton et al .1995 and *Campylobacter coli* and Chatelain 1973 are subjective synonyms, Int. J. Sys. Bacteriol., vol. 47, no. 4, 1055-1060.

Vandamme, P., Falsen, E., Rossau, R., Hoste, B., Segers, P., Tytgat, R., and De Ley, J. 1991, Revision of *Campylobacter, Helicobacter,* and *Wolinella* taxonomy: emendation of generic descriptions and proposal of *Arcobacter* gen. nov, Int. J. Syst. Bacteriol., vol. 41, no. 1, 88-103.

Vandamme, P. and On, S. L. 2001, Recommendations of the subcommittee on the taxonomy of *Campylobacter* and related bacteria, Int. J. Syst. Evol. Microbiol., vol. 51, no. 2, 719-721.

Vanniasinkam, T., Lanser, J. A., and Barton, M. D. 1999, PCR for the detection of *Campylobacter* spp. in clinical specimens, Lett. Appl. Microbiol., vol. 28, no. 1, 52-56.

Vargas, A. C., Costa, M. M., Vainstein, M. H., Kreutz, L. C., and Neves, J. P. 2003, Phenotypic and molecular characterization of bovine *Campylobacter fetus* strains isolated in Brazil, Vet. Microbiol., vol. 19;93, no. 2, 121-132.

Viscidi, R. P. and Demma, J. C. 2003, Genetic diversity of *Neisseria gonorrhoeae* housekeeping genes, J. Clin. Microbiol., vol. 41, no. 1, 197-204.

Volokhov, D., Chizhikov, V., Chumakov, K., and Rasooly, A. 2003, Microarray-based identification of thermophilic *Campylobacter jejuni, C. coli, C. lari, and C. upsaliensis,* J. Clin. Microbiol., vol. 41, no. 9, 4071-4080.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kulper, M., & Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acid Res., vol. 23, 4407-4411.

Waage, A. S., Verdund, T., Lund, V., and Kapperud, G. 1999, Detection of small numbers of *Campylobacter jejuni* and *Campylobacter coli* in environmental samples by a semi-nested PCR assay, Appl. Environ. Microbiol., vol. 65, no. 4, 1636-1643.

Waegel, A. and Nachamkin, I. 1996, Detection and molecular typing of *Campylobacter jejuni* in fecal samples by the Polymerase Chain Reaction, Mol. Cell. Probes, vol. 10, 75-80.

Wallace, J. S., Stanley, K. N., and Jones, K. 1998, The colonization of turkeys by thermophilic campylobacters, J. Appl. Microbiol., vol. 85, no. 2, 224-230.

Walters, J. J., Muhammad, W., Fox, K. F., Fox, A., Xie, D., Creek, K. E., and Pirisi, L. 2001, Genotyping single nucleotide polymorphisms using intact polymerase chain reaction products by electrospray quadrupole mass spectrometry, Rapid Commun. Mass Spectrom., vol. 15, no. 18, 1752-1759.

Wang, G., Clark, C. G., Taylor, T. M., Pucknell, C., Barton, C., Price, L., Woodward, D. L., and Rodgers, F. G. 2002, Colony multiplex PCR assay for identification and differentiation of *Campylobacter jejuni, C. coli, C. lari, C. upsaliensis, and C. fetus* subsp. *fetus*, J. Clin. Microbiol., vol. 40, no. 12, 4744-4747.

Wang, H. 2002, Rapid methods for detection and enumeration of *Campylobacter* spp. in foods, J. AOAC Int., vol. 85, no. 4, 996-999.

Wang, H., Farber, J. M., Malik, N., and Saunders, G. 1999, Improved PCR detection of *Campylobacter jejuni* from chicken rinses by a simple sample preparation procedure, Int. J. Food Microbiol., vol. 52, 39-45.

Ward, C. P., Fensom, A. H., and Green, P. M. 2000, Biallelic discrimination assays for the three common Ashkenazi Jewish mutations and a common non-Jewish mutation, in Tay-Sachs disease, using fluorogenic TaqMan probes, Genet. Test., vol. 4, no. 4, 351-358.

Wareing, D. R. A., Ure, R, Colles, F. M., Bolton, F. J., Fox, A. J., Maiden, M. C. J., and Dingle, K. E. 2003, Reference isolates for the clonal complexes of *Campylobacter jejuni*, Lett. Appl. Microbiol., vol. 36, 106-110.

Wassenaar, T. M. 1997, Toxin production by *Campylobacter* spp, Clin. Microbiol. Rev., vol. 10, no. 3, 466-476.

Wassenaar, T. M. and Blaser, M. J. 1999, Pathophysiology of *Campylobacter* infections, Microbes. Infect., no. 1, 1023-1033.

Wassenaar, T. M., Bleumink-Pluym, N. M., and van der Zeijst, B. A. 1991, Inactivation of *Campylobacter jejuni* flagellin genes by homologous recombination demonstrates that *fla*A but not *fla*B is required for invasion, EMBO J, vol. 10, no. 8, 2055-2061.

Wassenaar, T. M., Fry, B. N., Lastovica, A. J., Wagenaar, J. A., Coloe, P. J., and Duim, B. 2000, Genetic characterization of *Campylobacter jejuni* O:41 isolates in relation with Guillain-Barre syndrome, J. Clin. Microbiol., vol. 38, no. 2, 874-876.

Wassenaar, T. M., Fry, B. N., and van der Zeijst, B. A. 1995, Variation of the flagellin gene locus of *Campylobacter jejuni* by recombination and horizontal gene transfer, Microbiology, vol. 141, no. 1, 95-101.

Wassenaar, T. M., Geilhausen, B., and Newell, D. G. 1998, Evidence of genomic instability in *Campylobacter jejuni* isolated from poultry, Appl. Environ. Microbiol., vol. 64, 1816-1821.

Wassenaar, T. M., Newell, D.G. 2000, Genotyping of *Campylobacter* spp. Appl. Environ. Microbiol., vol. 66, no. 1, 1-9.

Wegmuller, B., Luthy, J., and Candrian, U. 1993, Direct polymerase chain reaction detection of *Campylobacter jejuni* and *Campylobacter Coli* in raw milk and dairy products, Appl. Environ. Microbiol., vol. 59, no. 7, 2161-2165.

Weissman, S. J., Moseley, S. L., Dykhuizen, D. E., and Sokurenko, E. V. 2003, Enterobacterial adhesins and the case for studying SNPs in bacteria, Trends Microbiol., vol. 11, no. 3, 115-117.

Wempe, J. M., Genigeorgis, C. A., Farver, T. B., and Yusufu, H. I. 1983, Prevalence of *Campylobacter jejuni* in two California chicken processing plants, Appl. Environ. Microbiol., vol. 45, no. 2, 355-359.

Wheeler, J. G., Sethi, D., Cowden, J. M., Wall, P. G., Rodrigues, L. C., Tompkins, D. S., Hudson, M. J., and Roderick, P. J. 1999, Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive, BMJ, vol. 318, no. 7190, 1046-1050.

Willis, W. L. and Murray, C. 1997, *Campylobacter jejuni* seasonal recovery observations of retail market broilers, Poult. Sci., vol. 76, no. 2, 314-317.

Wilson, K., Ausubel, F. M., Brent, R., & Kingston, R. E. 1987a, "Preparation and Analysis of DNA," in *Current Protocols in Molecular Biology*, John Wiley & sons, New York, p. 2.5A.3-2.5A4.

Wilson, K., Ausubel, F. M., Brent, R., & Kingston, R. E. 1987b, "Preparation of Genomic DNA from Bacteria," in *Current Protocols in Molecular Biology.*, John Wiley and Sons, New York, p. 2.4.1-2.4.2.

Winer, J. B. 2001, Guillain Barre syndrome, J. Clin. Path: Mol. Pathol., vol. 54, 381-385.

Winkler, M. A., Uher, J., and Cepa, S. 1999, Direct analysis and identification of *Helicobacter* and *Campylobacter species* by MALDI-TOF mass spectrometry, Anal. Chem., vol. 71, no. 16, 3416-3419.

Winters, D. K., O'Leary, A. E., and Slavik, M. F. 1997, Rapid PCR with nested primers for direct detection of *Campylobacter jejuni* in chicken washes, Mol. Cell. Probes, vol. 11, no. 4, 267-271.

Winters, D. K., O'Leary, A. E., and Slavik, M. F. 1998, Polymerase chain reaction for rapid detection of *Campylobacter jejuni* in artificially contaminated foods, Lett. Appl. Microbiol., vol. 27, no. 3, 163-167.

Winters, D. K. and Slavik, M. F. 1995, Evaluation of a PCR based assay for specific detection of *Campylobacter jejuni* in chicken washes, Mol. Cell. Probes., vol. 9, no. 5, 307-310.

Winters, D. K. and Slavik, M. F. 2000, Multiplex PCR detection of *Campylobacter jejuni* and *Arcobacter butzleri* in food products, Mol. Cell. Probes., vol. 14, 95-99.

Woo, P. C., Leung, K. W., Tsoi, H. W., Wong, S. S., Teng, J. L., and Yuen, K. Y. 2002, Thermo-tolerant *Campylobacter fetus* bacteraemia identified by 16S ribosomal RNA gene sequencing: an emerging pathogen in immunocompromised patients, J. Med. Microbiol., vol. 51, no. 9, 740-746.

Wooldridge, K. G., Williams, P. H., and Ketley, J. M. 1996, Host signal transduction and endocytosis of *Campylobacter jejuni*, Microb. Pathog., vol. 21, no. 4, 299-305.

Wu, T. L., Su, L. H., Chia, J. H., Kao, T. M., Chiu, C. H., Kuo, A. J., and Sun, C. F. 2002, Molecular epidemiology of nalidixic acid-resistant campylobacter isolates from humans and poultry by pulsed-field gel electrophoresis and flagellin gene analysis, Epidemiol. Infect., vol. 129, no. 1, 227-231.

Yan, W., Chang, N., and Taylor, D. E. 1991, Pulsed Field Gel Electrophoresis of *Campylobacter jejuni* and *Campylobacter coli* genomic DNA and its epidemiologic application, J. Infect. Dis., vol. 163, 1068-1072.

Yang, C., Jiang, Y., Huang, K., Zhu, C., and Yin, Y. 2003, Application of real-time PCR for quantitative detection of *Campylobacter jejuni* in poultry, milk and environmental water, FEMS Immunol. Med. Microbiol., vol. 38, no. 3, 265-271.

Yao, R., Burr, D. H., Doig, P., Trust, T. J., Niu, H., and Guerry, P. 1994, Isolation of motile and non-motile insertional mutants of *Campylobacter jejuni*: the role of motility in adherence and invasion of eukaryotic cells, Mol. Microbiol., vol. 14, no. 5, 883-893.

Ye, F., Li, M. S., Taylor, J. D., Nguyen, Q., Colton, H. M., Casey, W. M., Wagner, M., Weiner, M. P., and Chen, J. 2001, Fluorescent microsphere-based readout technology for multiplexed human single nucleotide polymorphism analysis and bacterial identification, Hum. Mutat., vol. 17, no. 4, 305-316.

Young, V. B., Dangler, C. A., Fox, J. G., and Schauer, D. B. 2000, Chronic atrophic gastritis in SCID mice experimentally infected with *Campylobacter fetus*, Infect. Immun., vol. 68, no. 4, 2110-2118.

Yuki, N. 2001, Infectious origins of, and molecular mimicry in, Guillain-Barre and Fisher syndromes, Lancet Infect. Dis., vol. 1, no. 1, 29-37.

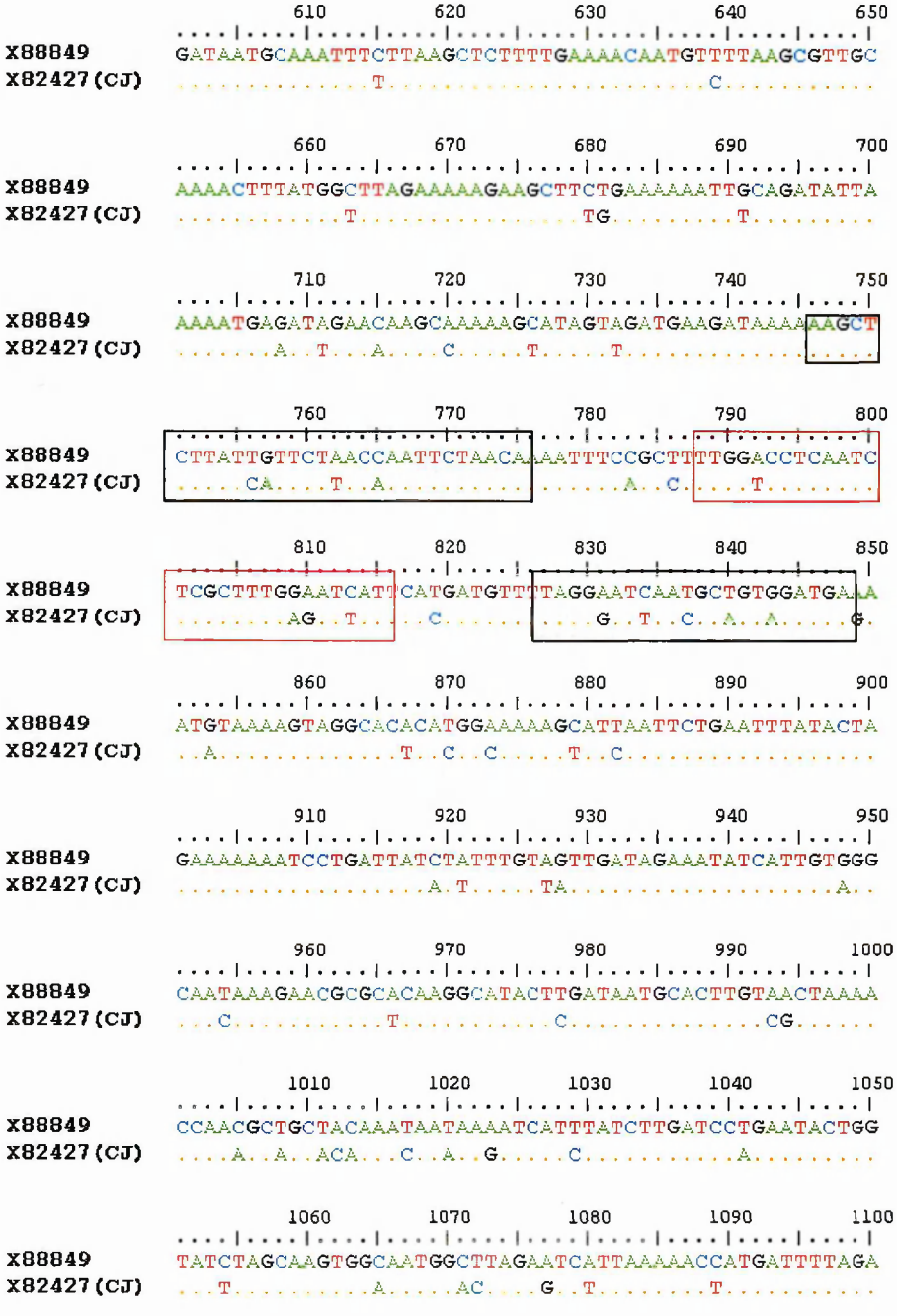
Zhao, Y., Yu, M., Miller, J. W., Chen, M., Bremer, E. G., Kabat, W., and Yogev, R. 2002, Quantification of HIV proviral load by using Taqman technology, J. Clin. Microbiol., vol. 40, no. 2, 675-678.

Zimmer, M., Barnhart, H., Idris, U., and Lee, M. D. 2003, Detection of *Campylobacter jejuni* strains in the water lines of a commercial broiler house and their relationship to the strains that colonized the chickens, Avian Dis., vol. 47, no. 1, 101-107.

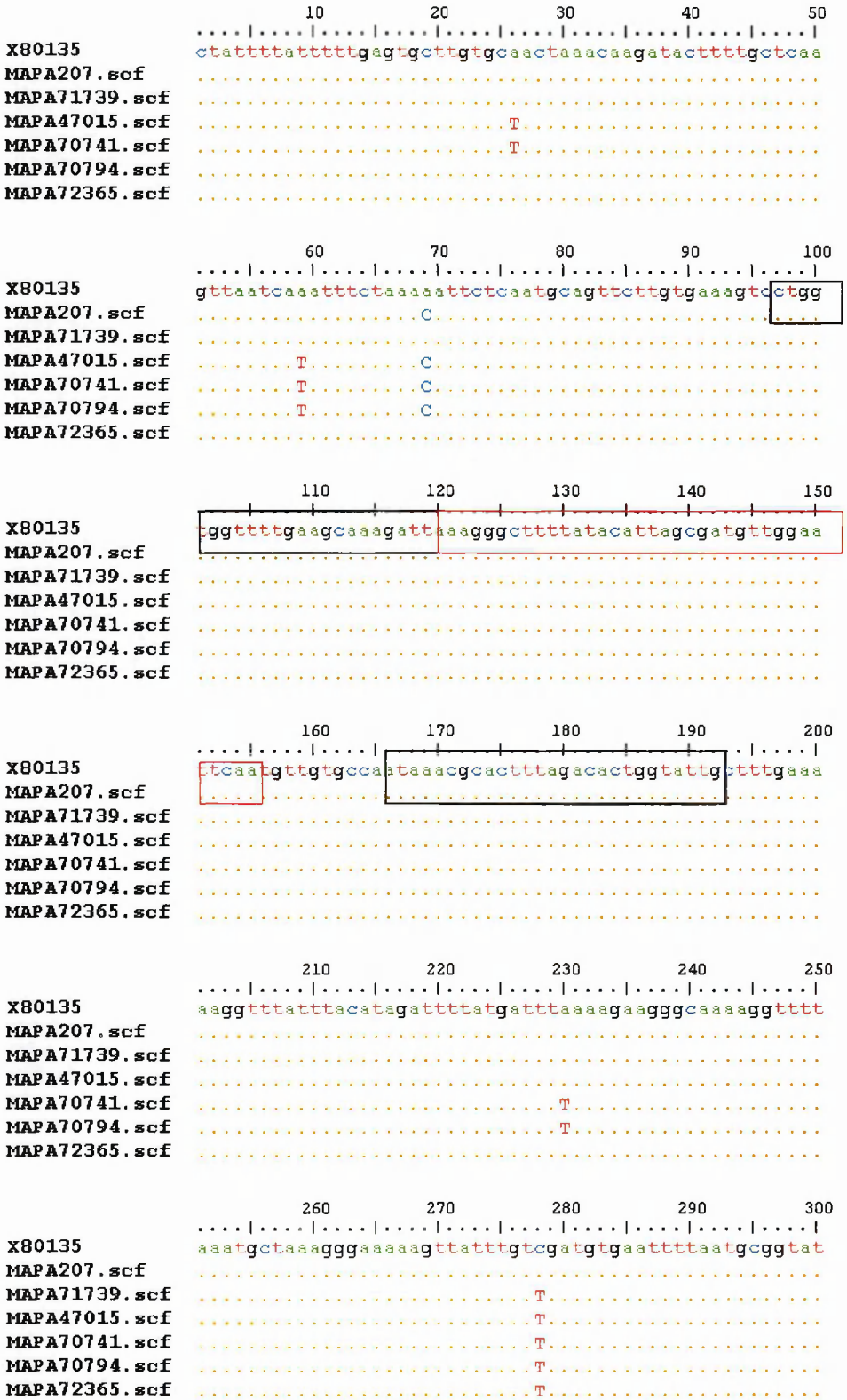
Ziprin, R. L., Young, C. R., Stanker, L. H., Hume, M. E., and Konkel, M. E. 1999, The absence of cecal colonization of chicks by a mutant of *Campylobacter jejuni* not expressing bacterial fibronectin-binding protein, Avian Dis., vol. 43, no. 3, 586-589.

Appendix 1.

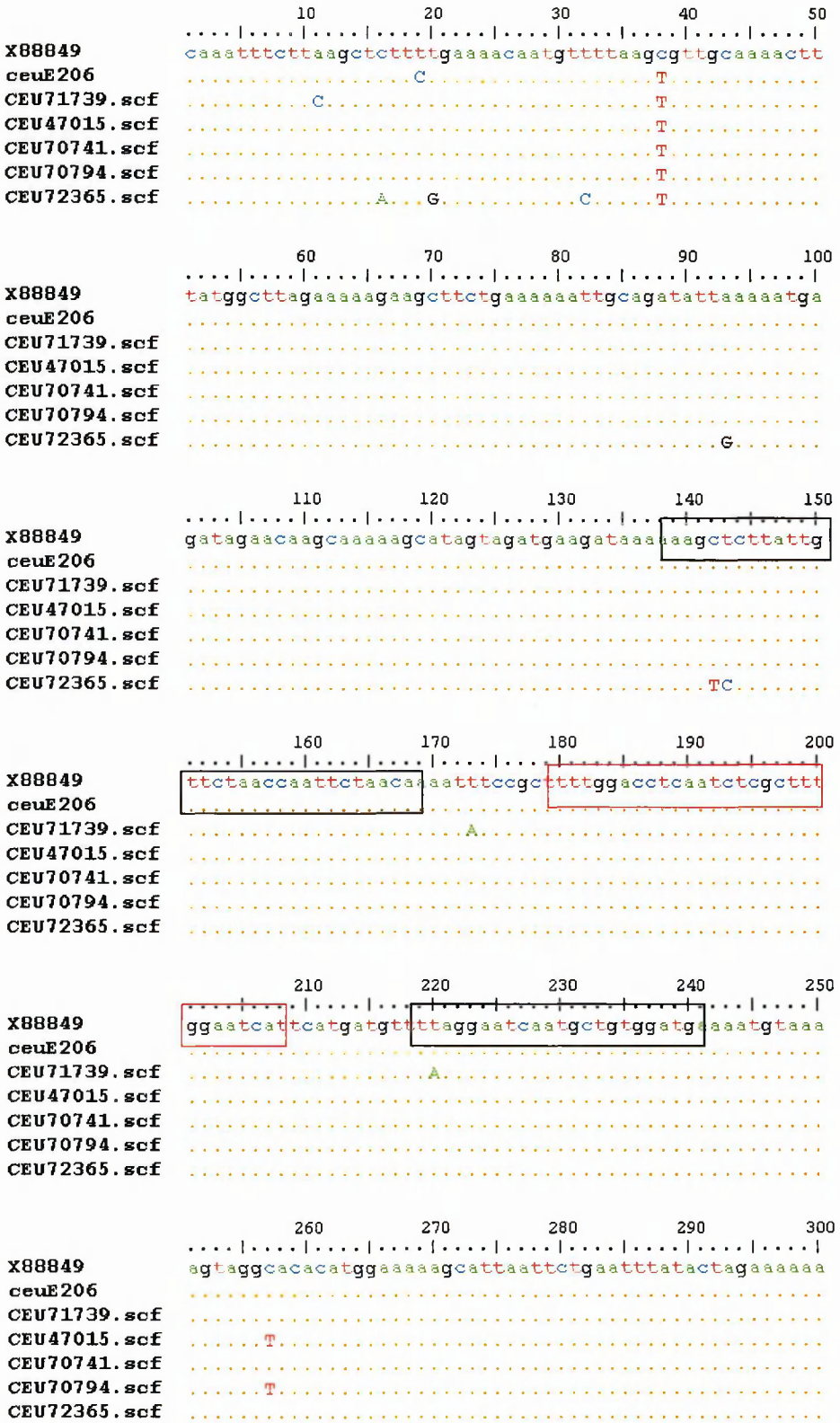
Appendix 1.1 Partial alignment of the *ceu*E gene from *C.coli* (X88849) and equivalent gene from *C.jejuni* (X82427) showing variation within the probe and primer regions. (Primer regions indicated by black rectangles, probe region indicated by red rectangle).



Appendix 1.2 Alignment of *map*A sequences from mixed isolates, using published sequence X80135 as the consensus. Included as a control is the known *C.jejuni* strain 207. Position of *map*A primers (black rectangle) position of probe (red rectangle)

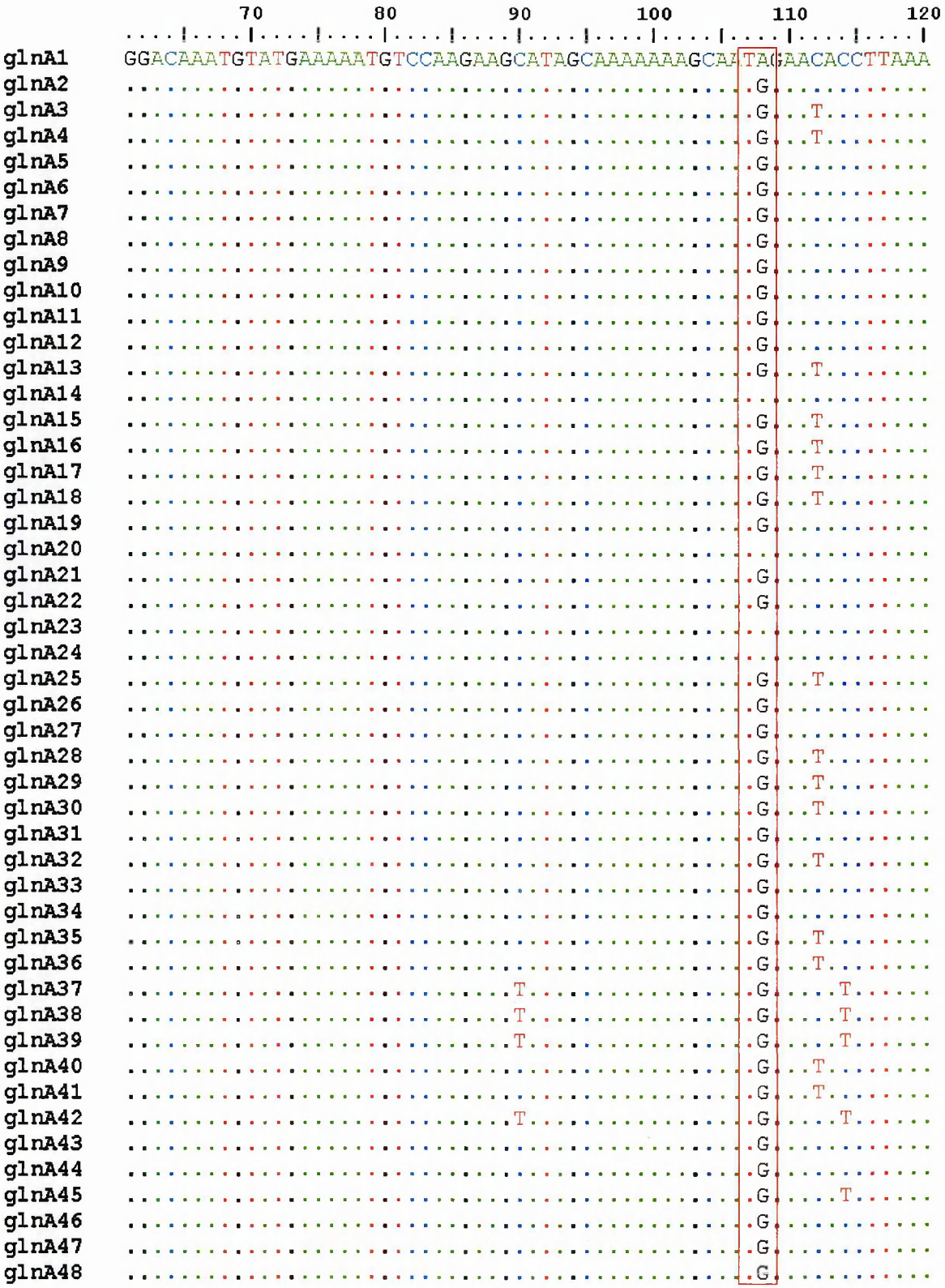


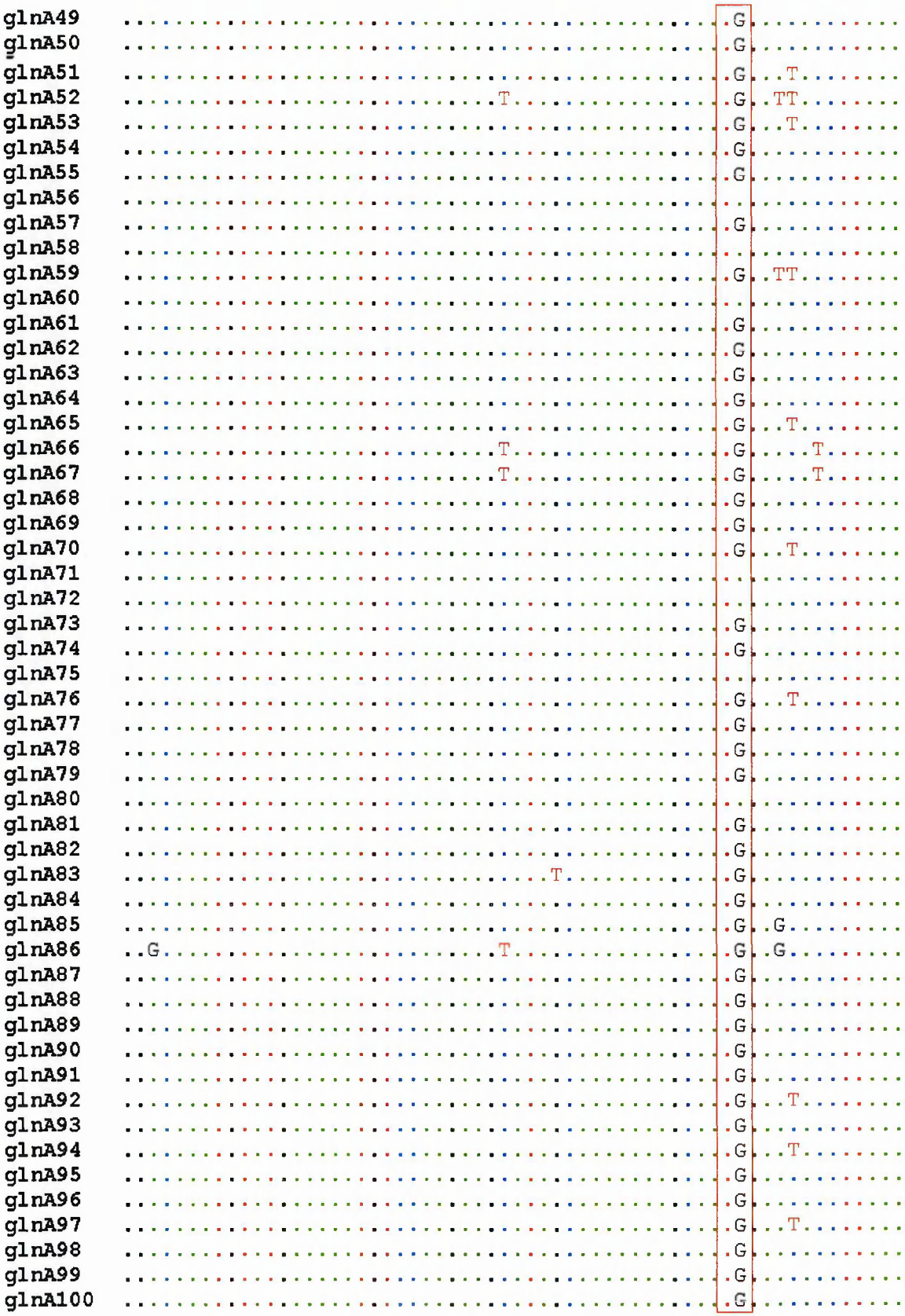
Appendix 1.3 Alignment of *ceu*E sequences from mixed isolates, using published sequence X88849 as the consensus. Included as a control is the known *C.coli* strain 206. Position of *ceuE* primers (black rectangle) position of probe (red rectangle)

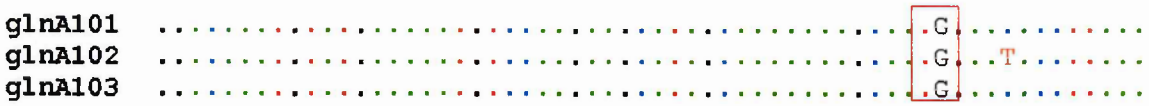


Appendix 2.

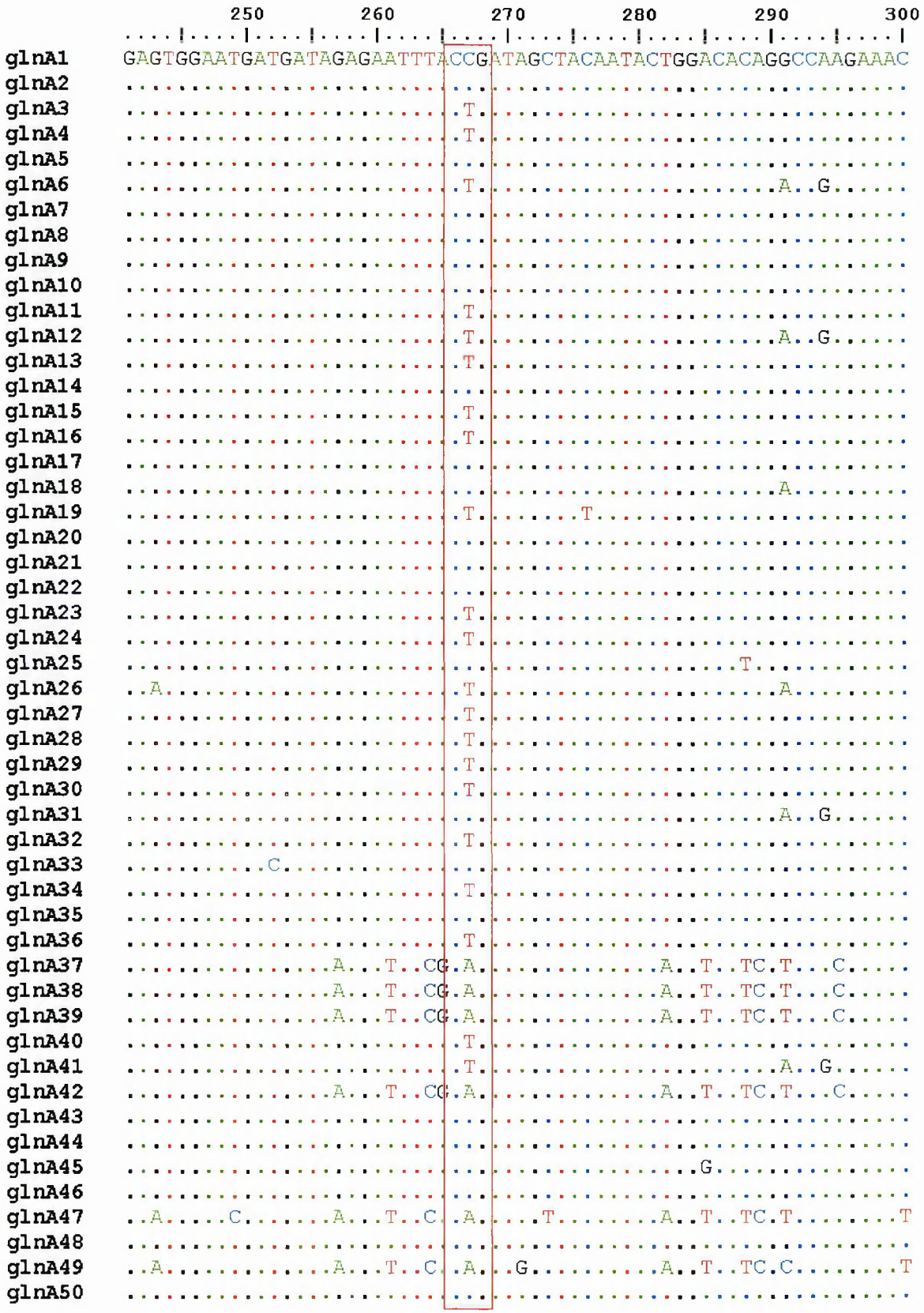
Appendix 2.1 *gln*A alignment showing SNP at 108bp on *gln*A1

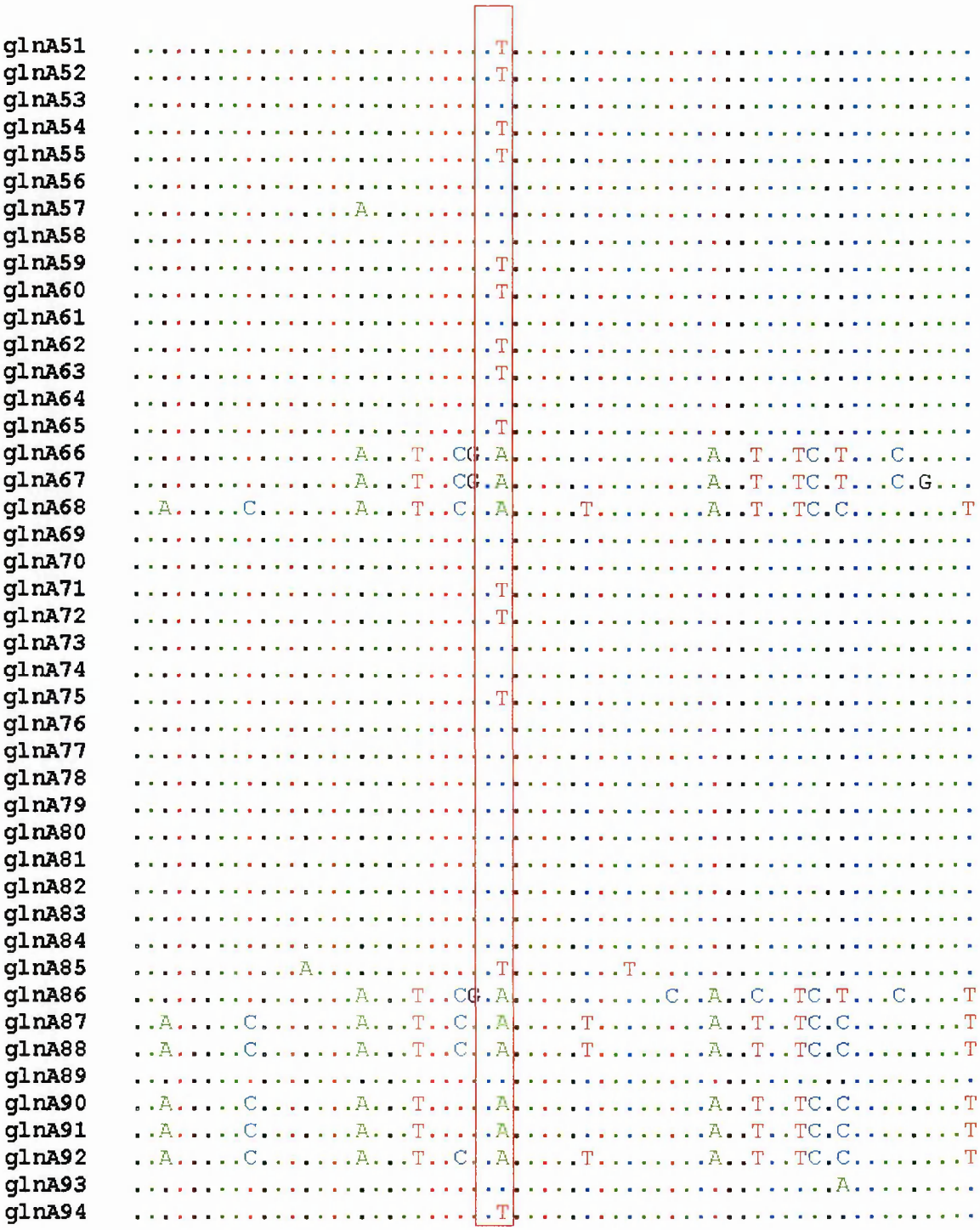


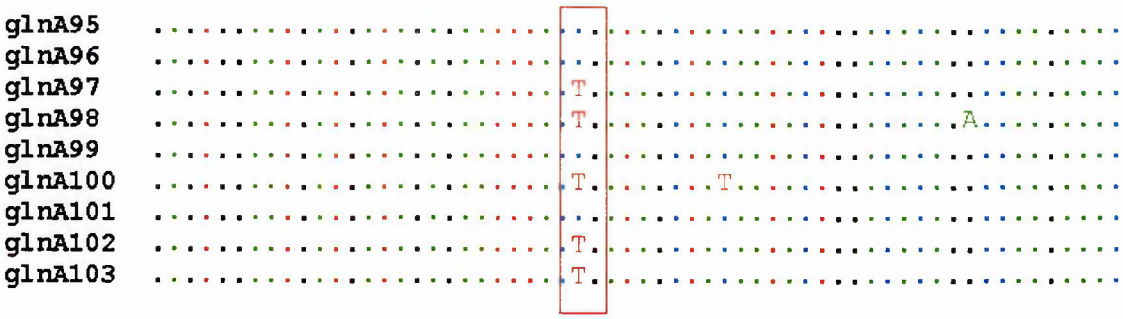

glnA49
glnA50
glnA51
glnA52
glnA53
glnA54
glnA55
glnA56
glnA57
glnA58
glnA59
glnA60
glnA61
glnA62
glnA63
glnA64
glnA65
glnA66
glnA67
glnA68
glnA69
glnA70
glnA71
glnA72
glnA73
glnA74
glnA75
glnA76
glnA77
glnA78
glnA79
glnA80
glnA81
glnA82
glnA83
glnA84
glnA85
glnA86
glnA87
glnA88
glnA89
glnA90
glnA91
glnA92
glnA93
glnA94
glnA95
glnA96
glnA97
glnA98
glnA99
glnA100


glnA101
glnA102
glnA103

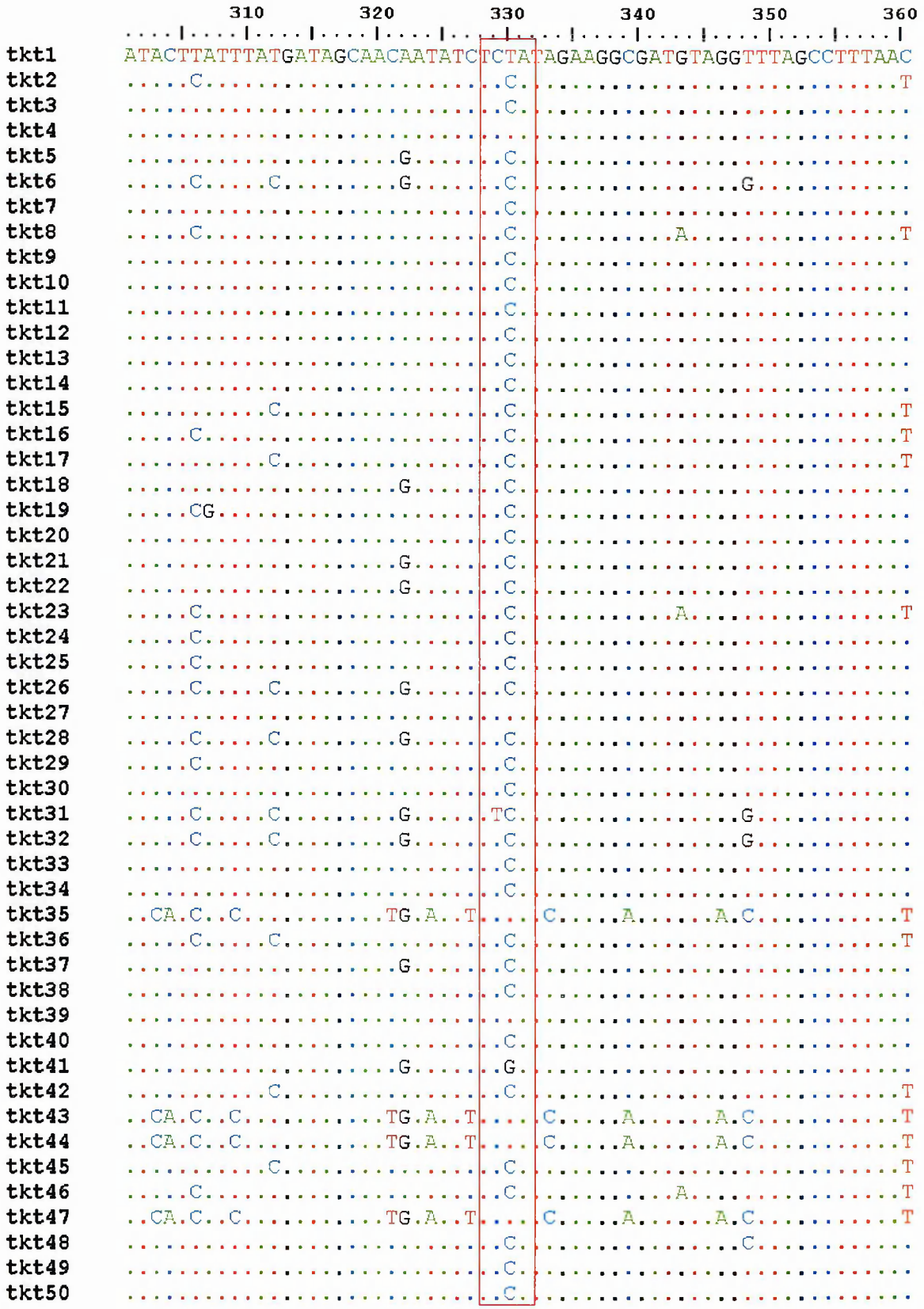
Appendix 2.2 *gln*A alignment showing SNP at 267bp on *gln*A1

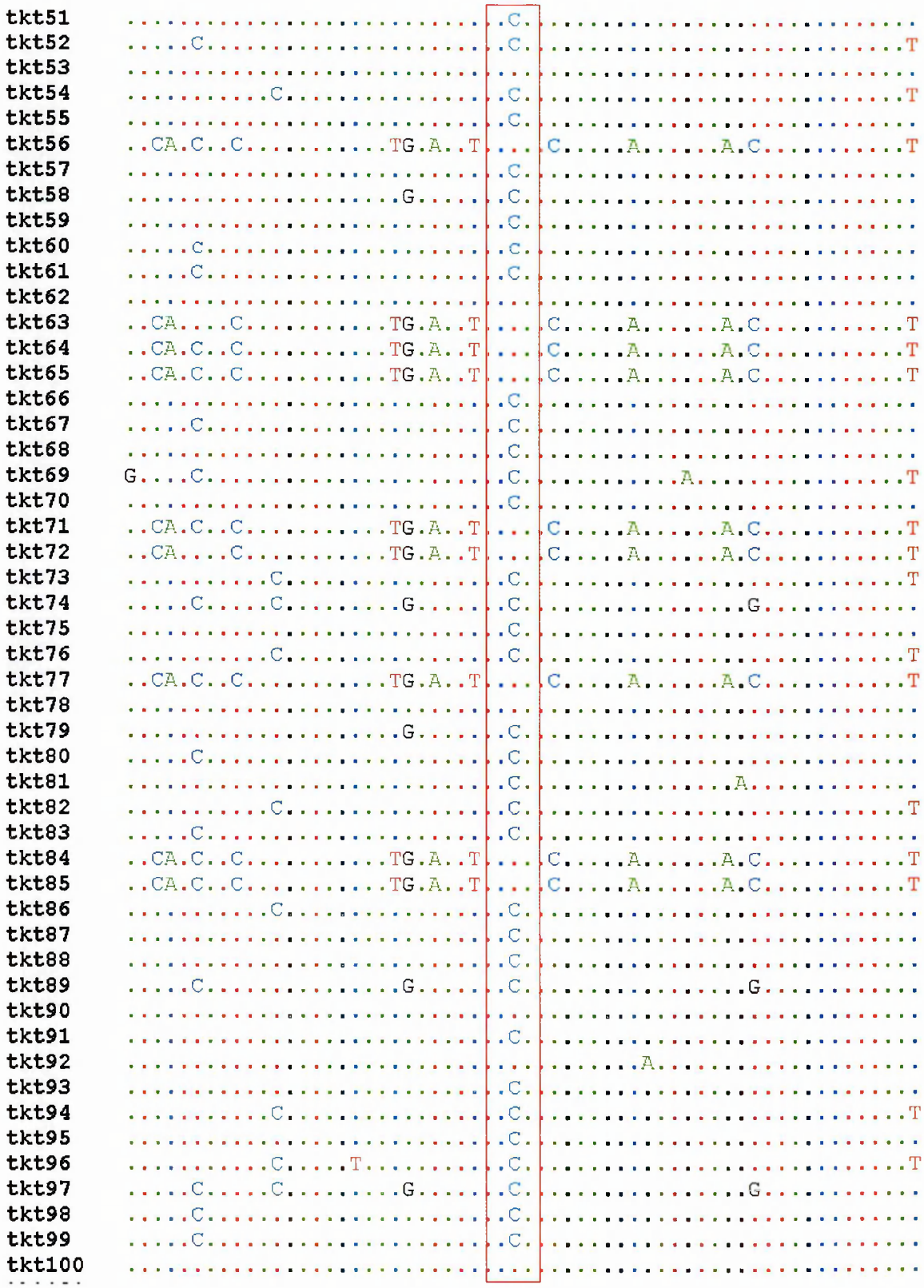


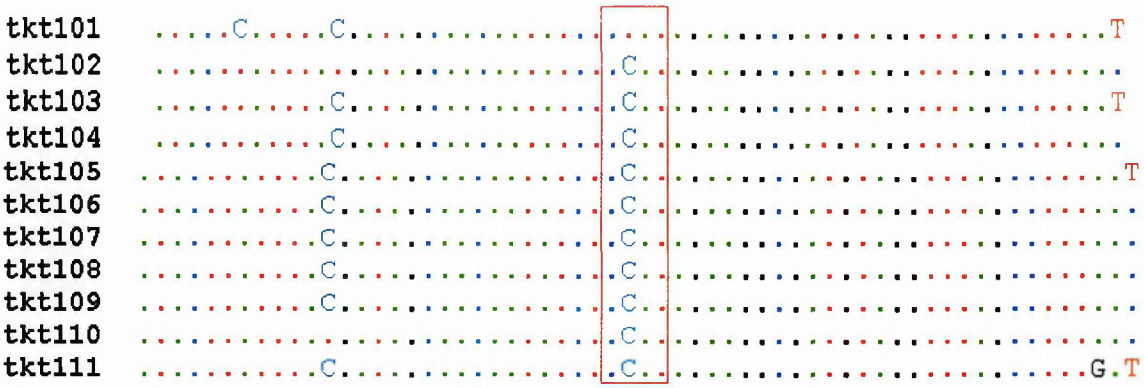

glnA51
glnA52
glnA53
glnA54
glnA55
glnA56
glnA57
glnA58
glnA59
glnA60
glnA61
glnA62
glnA63
glnA64
glnA65
glnA66
glnA67
glnA68
glnA69
glnA70
glnA71
glnA72
glnA73
glnA74
glnA75
glnA76
glnA77
glnA78
glnA79
glnA80
glnA81
glnA82
glnA83
glnA84
glnA85
glnA86
glnA87
glnA88
glnA89
glnA90
glnA91
glnA92
glnA93
glnA94


glnA95
glnA96
glnA97
glnA98
glnA99
glnA100
glnA101
glnA102
glnA103

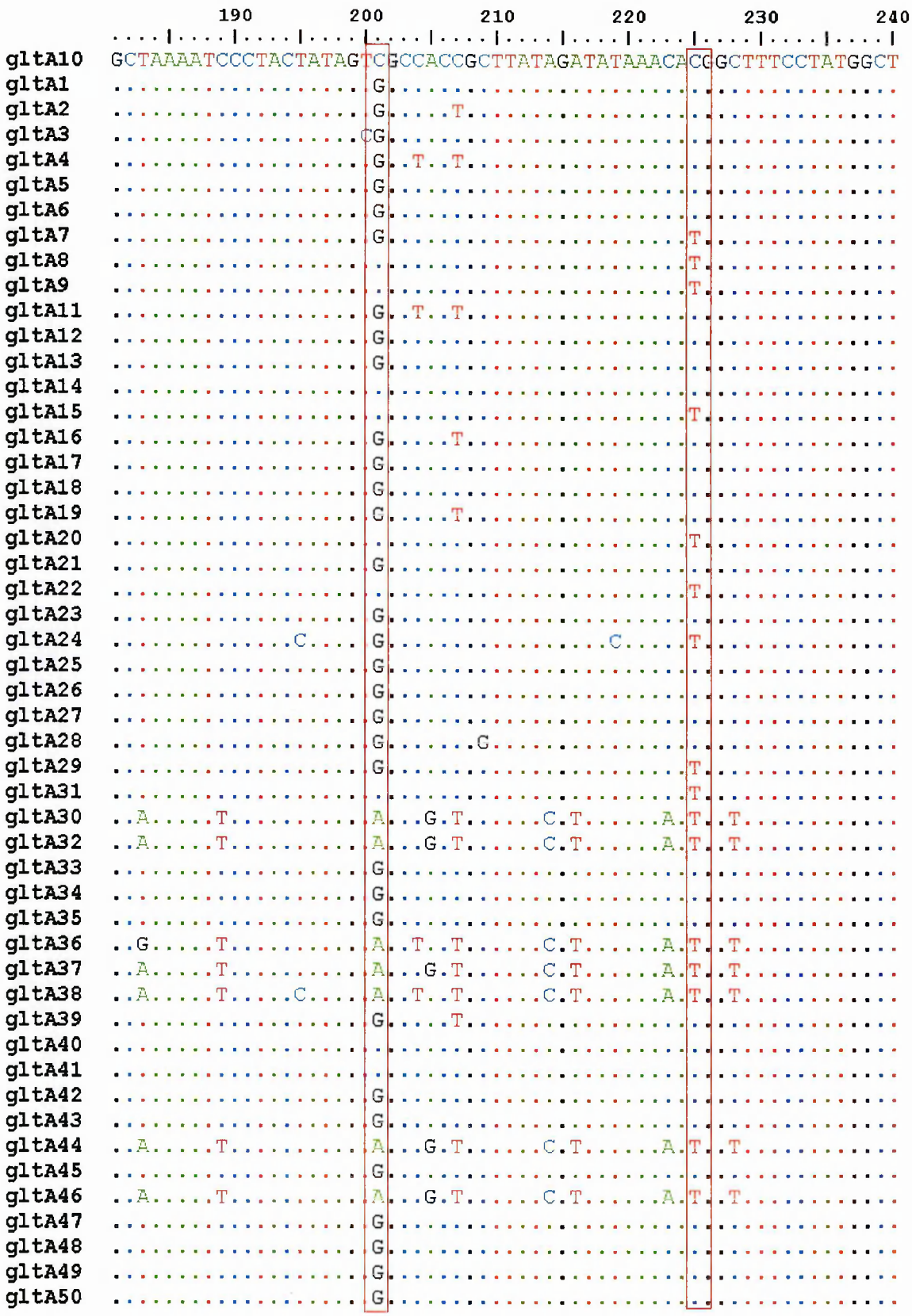
Appendix 2.3 *tkt* alignment showing SNP at 330bp on *tkt* 1

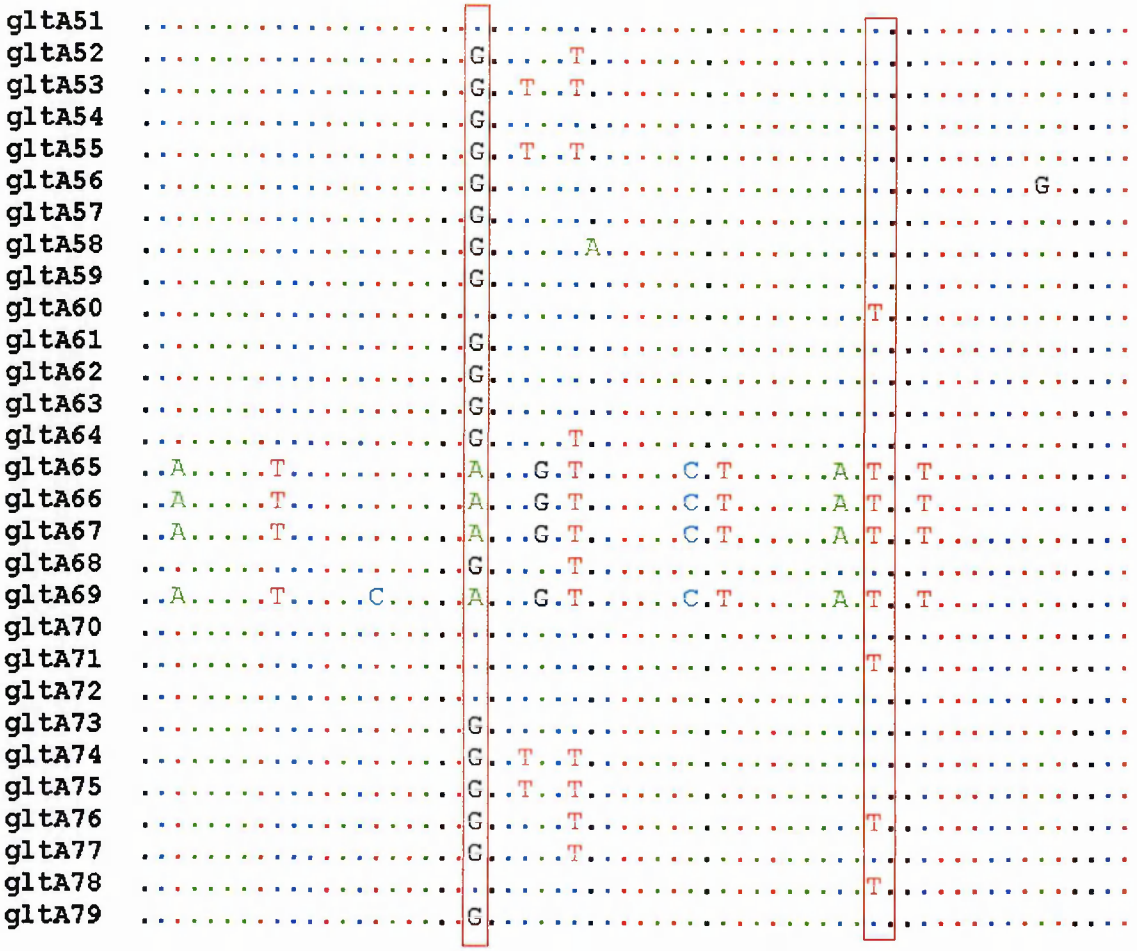



tkt51
tkt52
tkt53
tkt54
tkt55
tkt56
tkt57
tkt58
tkt59
tkt60
tkt61
tkt62
tkt63
tkt64
tkt65
tkt66
tkt67
tkt68
tkt69
tkt70
tkt71
tkt72
tkt73
tkt74
tkt75
tkt76
tkt77
tkt78
tkt79
tkt80
tkt81
tkt82
tkt83
tkt84
tkt85
tkt86
tkt87
tkt88
tkt89
tkt90
tkt91
tkt92
tkt93
tkt94
tkt95
tkt96
tkt97
tkt98
tkt99
tkt100

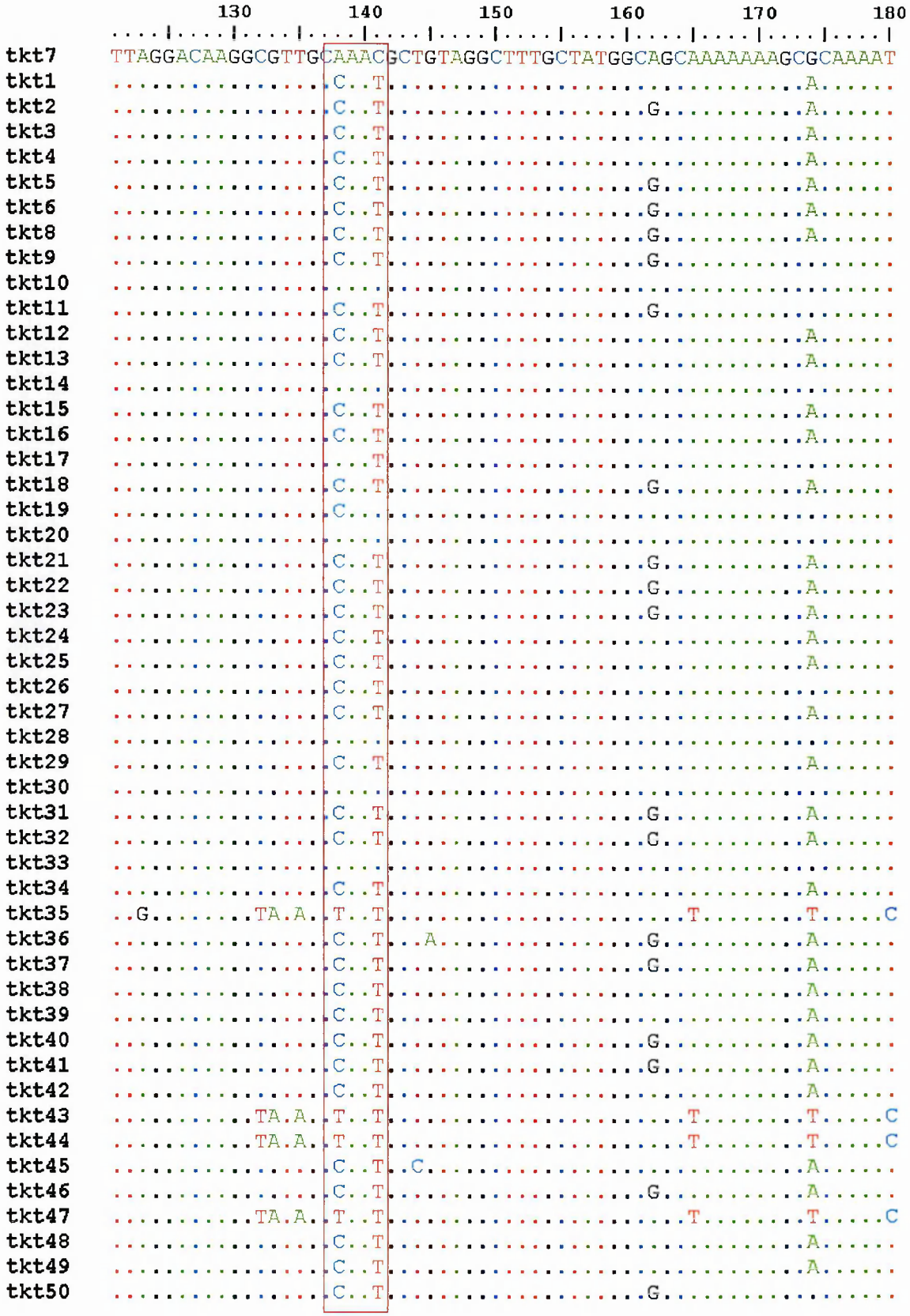

tkt101
tkt102
tkt103
tkt104
tkt105
tkt106
tkt107
tkt108
tkt109
tkt110
tkt111

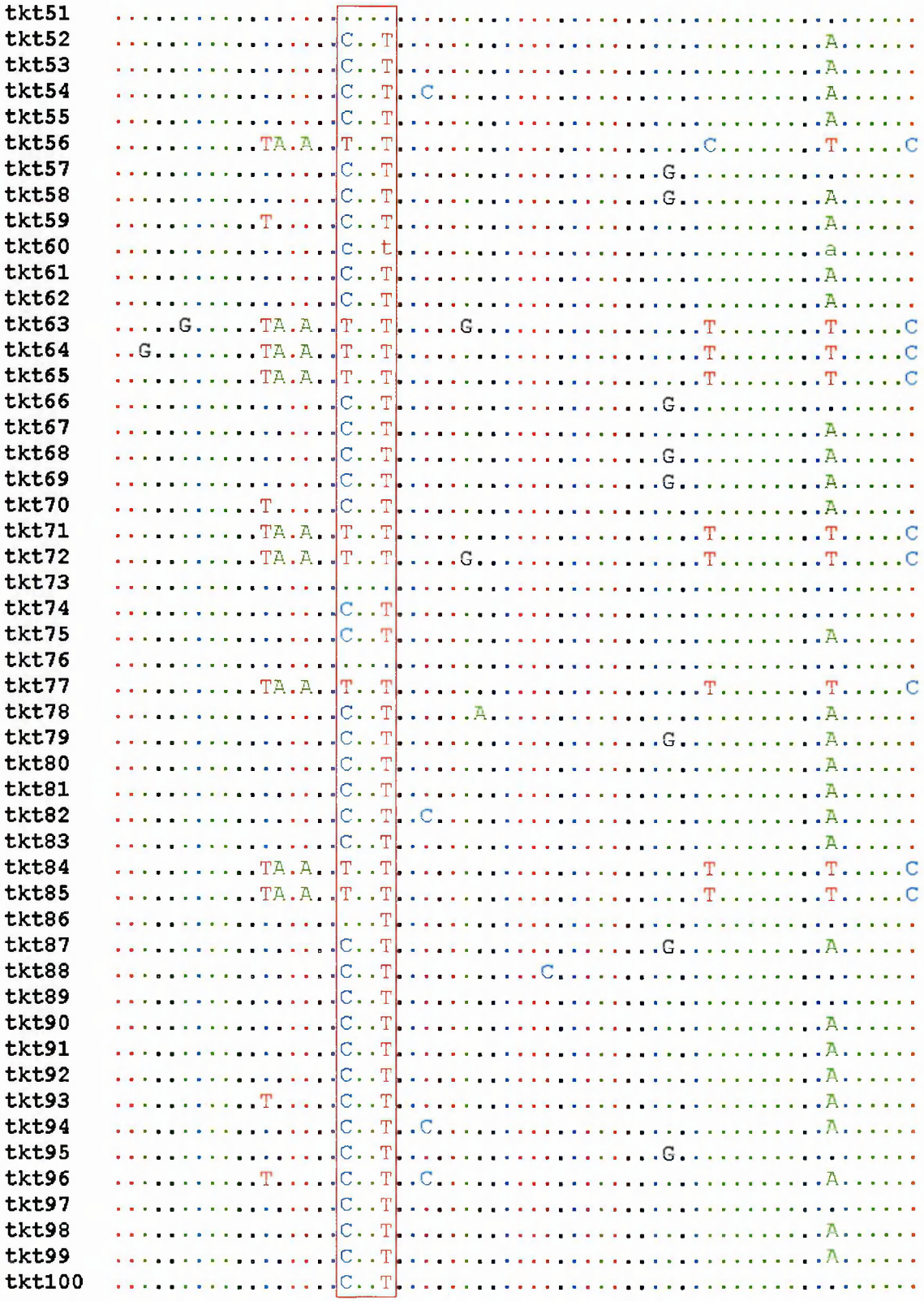
Appendix 2.4 *glt*A alignment showing SNPs at 201and 225bp on *glt*A10

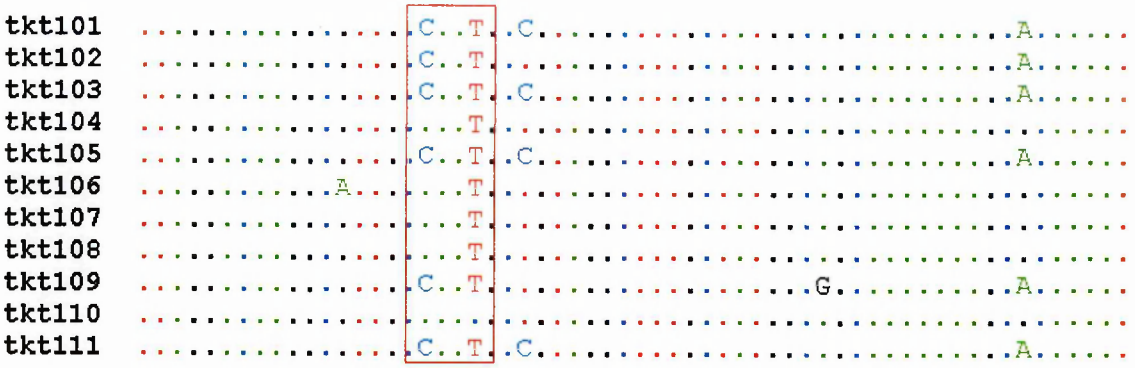



gltA51
gltA52
gltA53
gltA54
gltA55
gltA56
gltA57
gltA58
gltA59
gltA60
gltA61
gltA62
gltA63
gltA64
gltA65
gltA66
gltA67
gltA68
gltA69
gltA70
gltA71
gltA72
gltA73
gltA74
gltA75
gltA76
gltA77
gltA78
gltA79

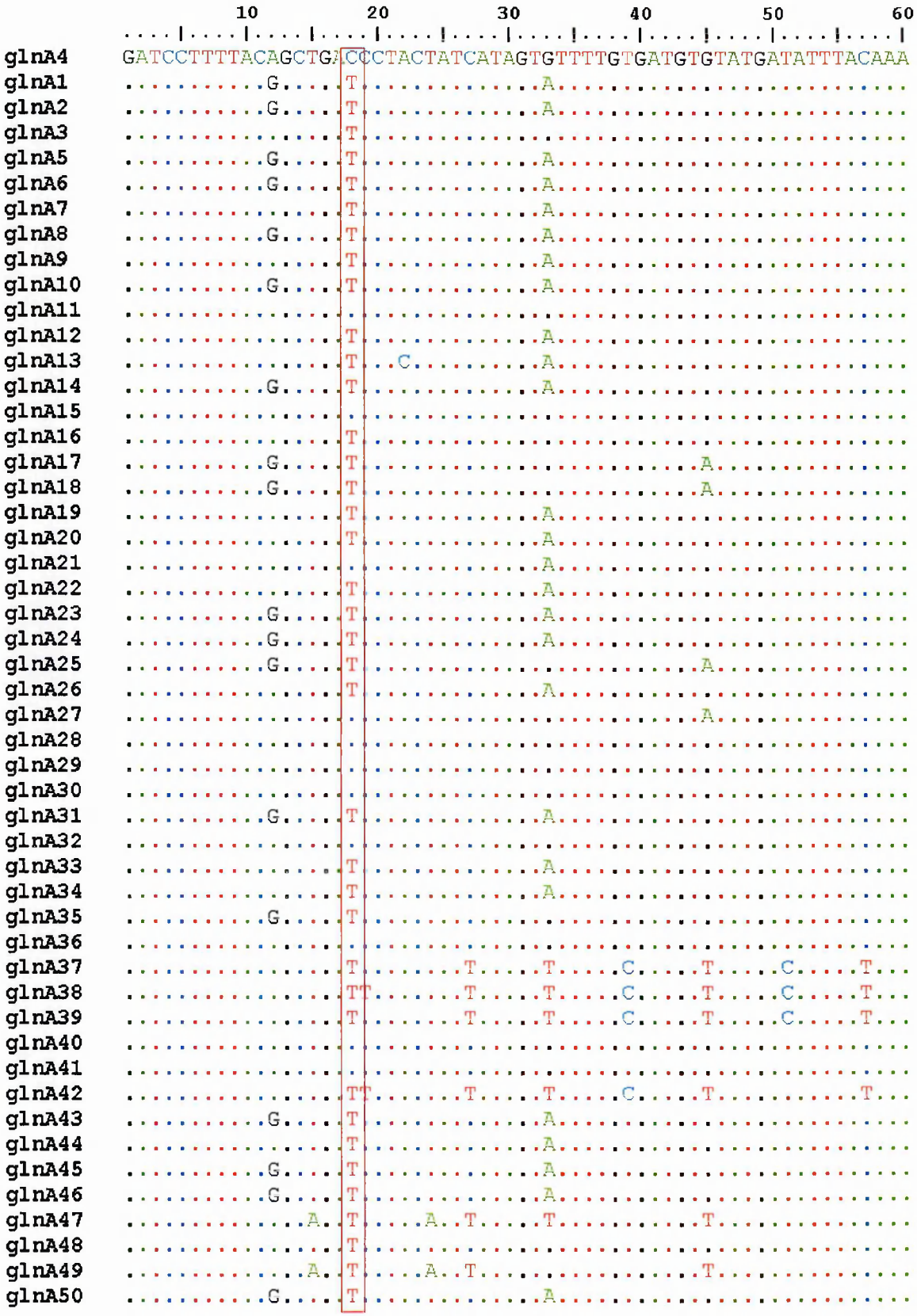
Appendix 2.5 ***tkt_*** **alignment showing SNPs at 138 and 141bp on** ***tkt_7***

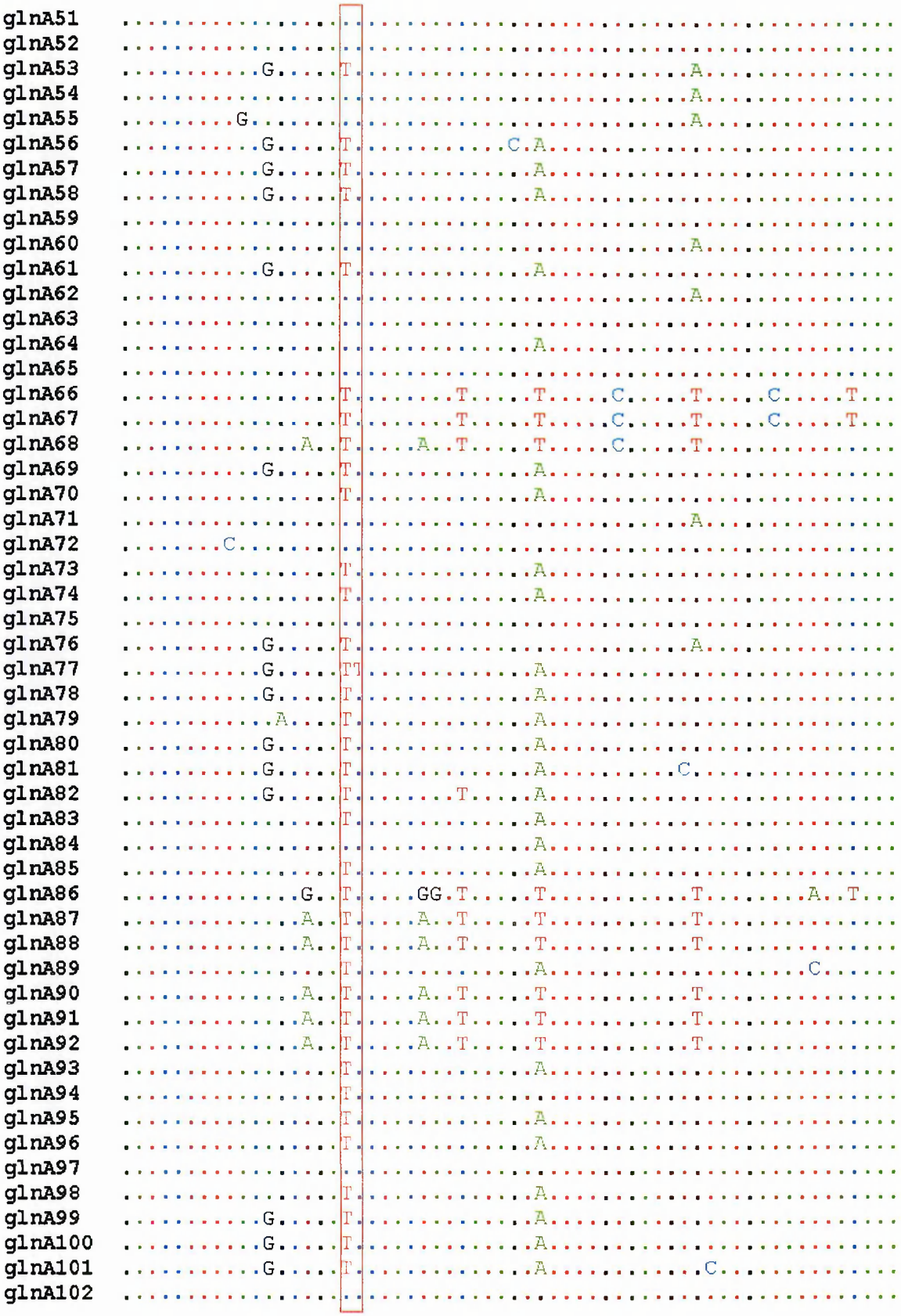



tkt51
tkt52
tkt53
tkt54
tkt55
tkt56
tkt57
tkt58
tkt59
tkt60
tkt61
tkt62
tkt63
tkt64
tkt65
tkt66
tkt67
tkt68
tkt69
tkt70
tkt71
tkt72
tkt73
tkt74
tkt75
tkt76
tkt77
tkt78
tkt79
tkt80
tkt81
tkt82
tkt83
tkt84
tkt85
tkt86
tkt87
tkt88
tkt89
tkt90
tkt91
tkt92
tkt93
tkt94
tkt95
tkt96
tkt97
tkt98
tkt99
tkt100


tkt101
tkt102
tkt103
tkt104
tkt105
tkt106
tkt107
tkt108
tkt109
tkt110
tkt111

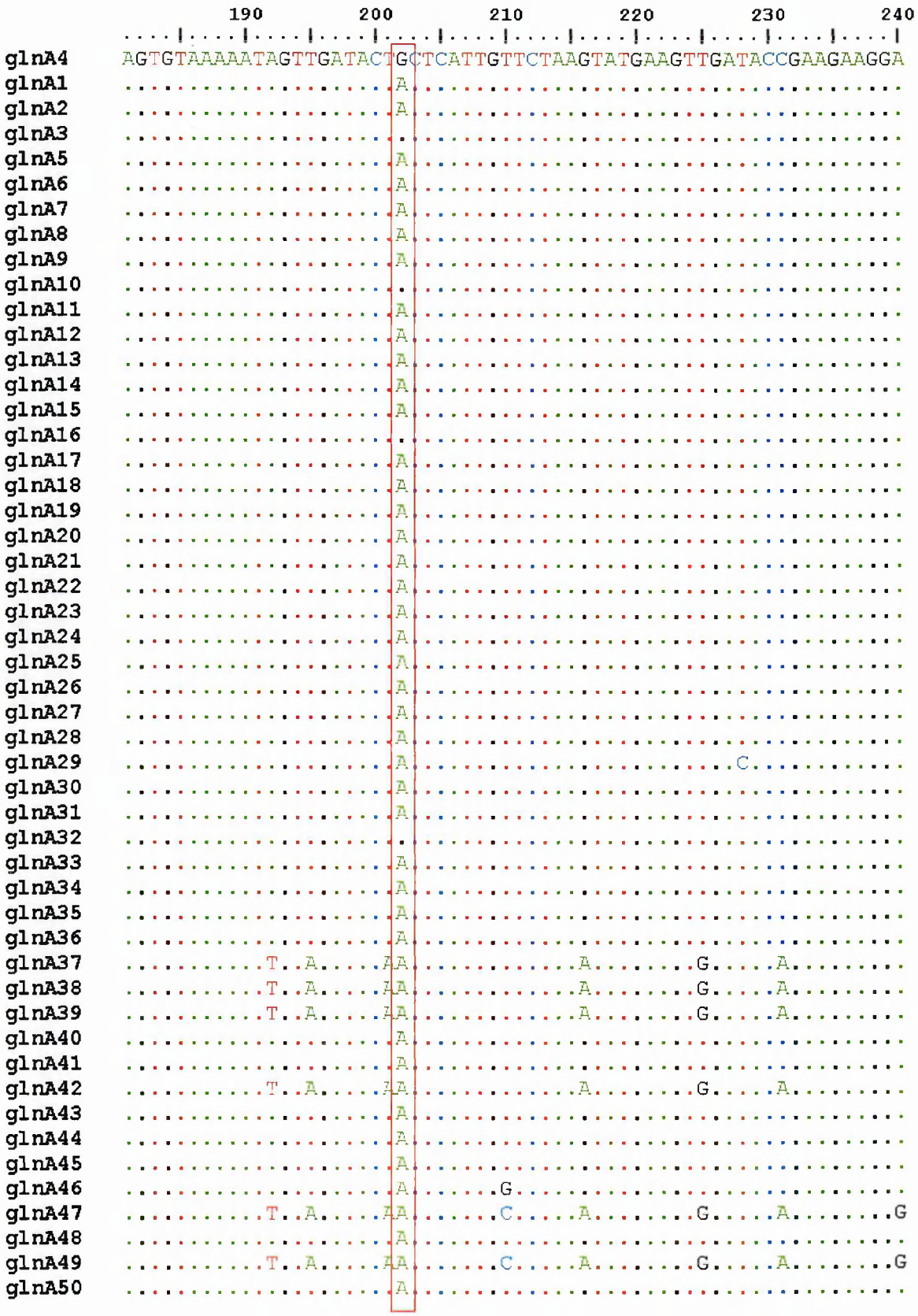
Appendix 2.6 *gln*A alignment showing SNP at 18bp on *gln*A4

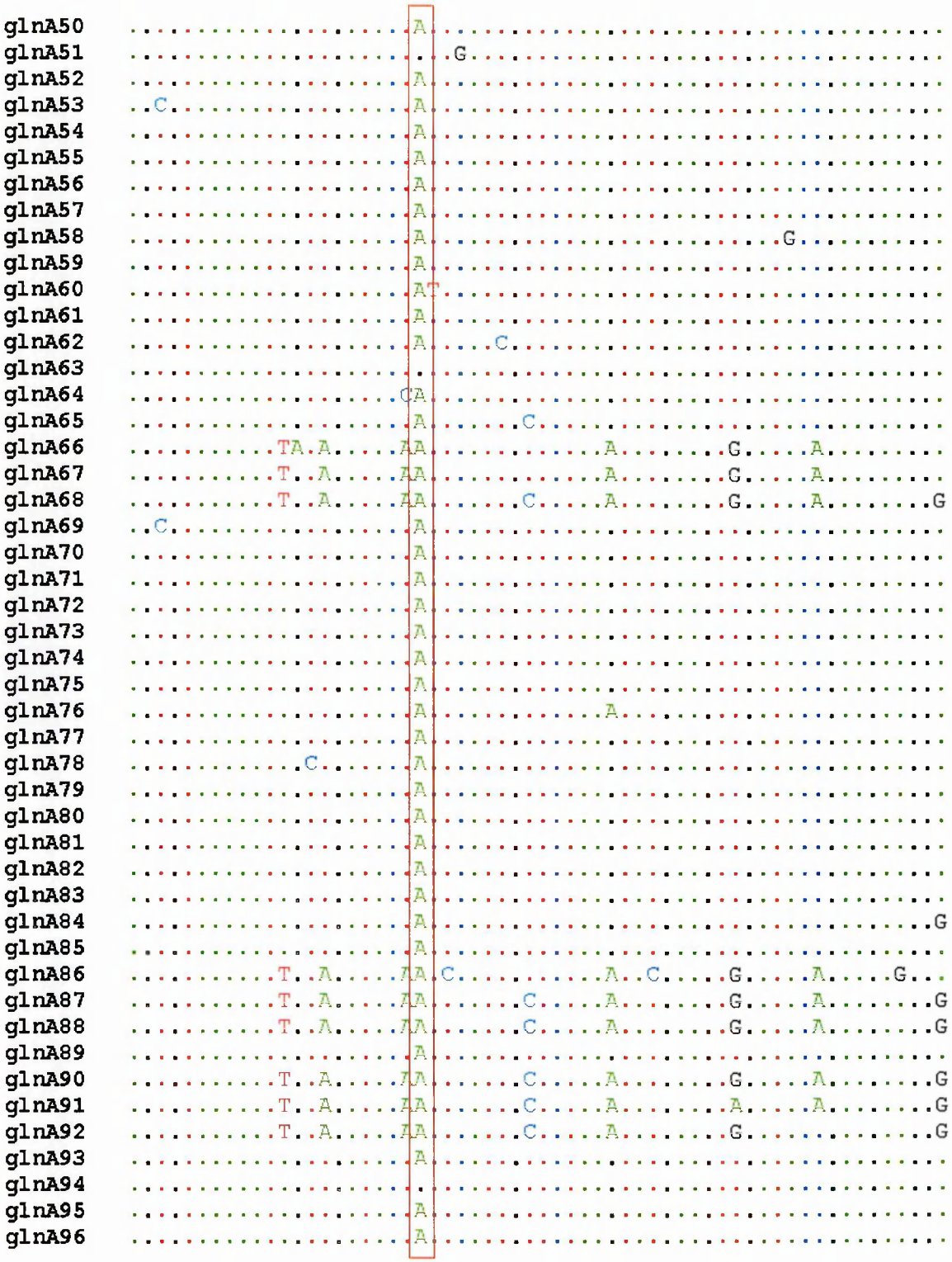


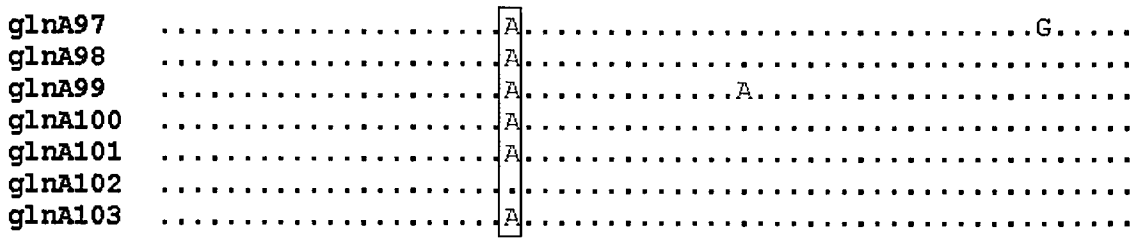

glnA51
glnA52
glnA53
glnA54
glnA55
glnA56
glnA57
glnA58
glnA59
glnA60
glnA61
glnA62
glnA63
glnA64
glnA65
glnA66
glnA67
glnA68
glnA69
glnA70
glnA71
glnA72
glnA73
glnA74
glnA75
glnA76
glnA77
glnA78
glnA79
glnA80
glnA81
glnA82
glnA83
glnA84
glnA85
glnA86
glnA87
glnA88
glnA89
glnA90
glnA91
glnA92
glnA93
glnA94
glnA95
glnA96
glnA97
glnA98
glnA99
glnA100
glnA101
glnA102


glnA103 ..........G....T.............A.....

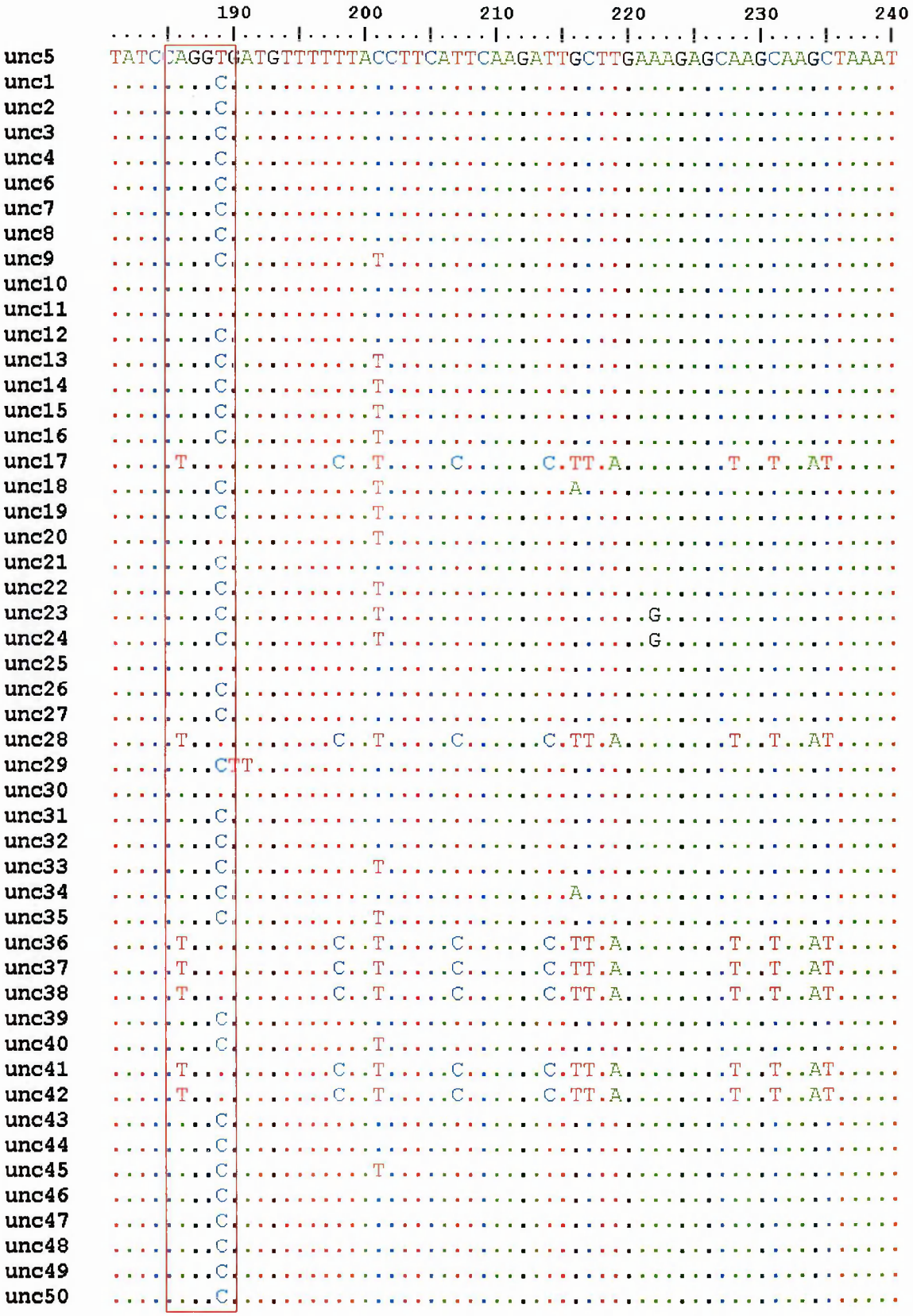
Appendix 2.7 *gln*A alignment showing SNP at 202bp on *gln*A4

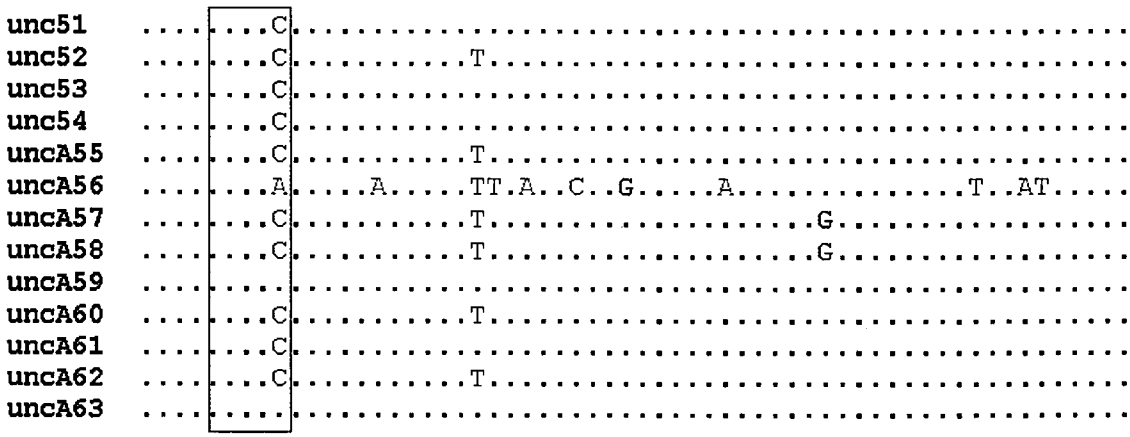



glnA50
glnA51
glnA52
glnA53
glnA54
glnA55
glnA56
glnA57
glnA58
glnA59
glnA60
glnA61
glnA62
glnA63
glnA64
glnA65
glnA66
glnA67
glnA68
glnA69
glnA70
glnA71
glnA72
glnA73
glnA74
glnA75
glnA76
glnA77
glnA78
glnA79
glnA80
glnA81
glnA82
glnA83
glnA84
glnA85
glnA86
glnA87
glnA88
glnA89
glnA90
glnA91
glnA92
glnA93
glnA94
glnA95
glnA96

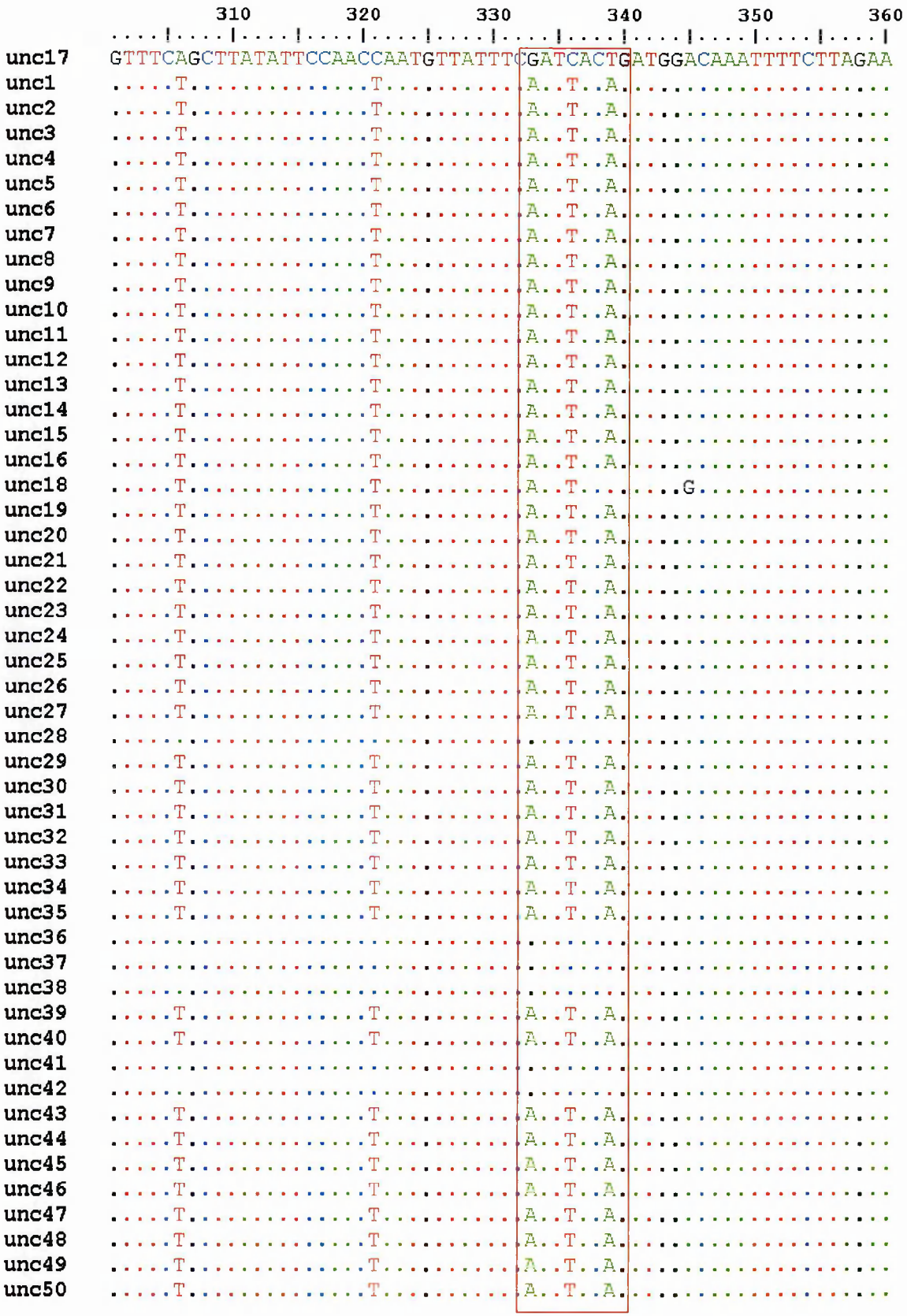

glnA97
glnA98
glnA99
glnA100
glnA101
glnA102
glnA103

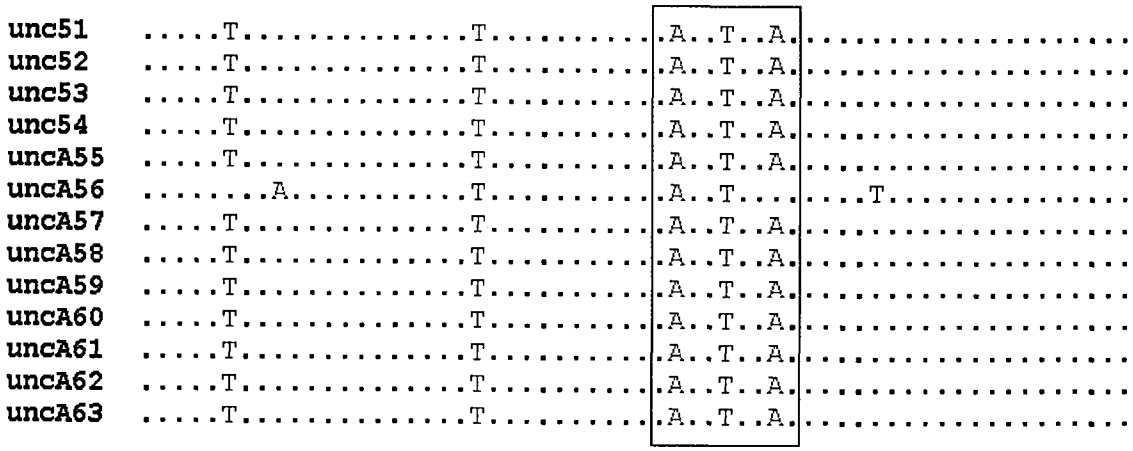
Appendix 2.8 *unc*A alignment showing SNPs at 186 and 189bp on *unc*A5



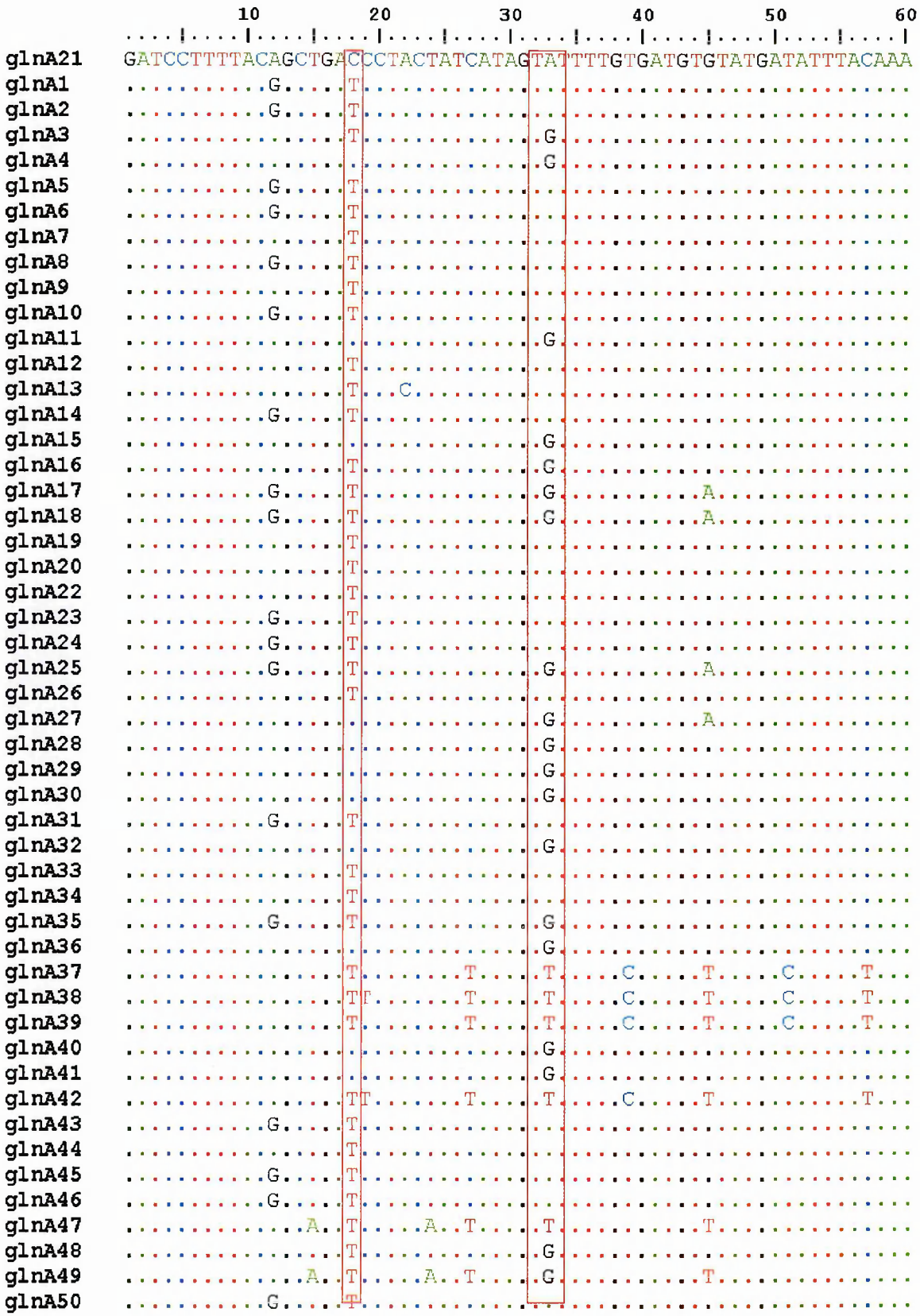

unc51 ....C.....
unc52 ....C.....T.....
unc53 ....C.....
unc54 ....C.....
uncA55 ....C.....T.....
uncA56 ....A.....A.....TT.A..C..G.....A.....T..AT.....
uncA57 ....C.....T.....G.....
uncA58 ....C.....T.....G.....
uncA59 .....
uncA60 ....C.....T.....
uncA61 ....C.....
uncA62 ....C.....T.....
uncA63 .....

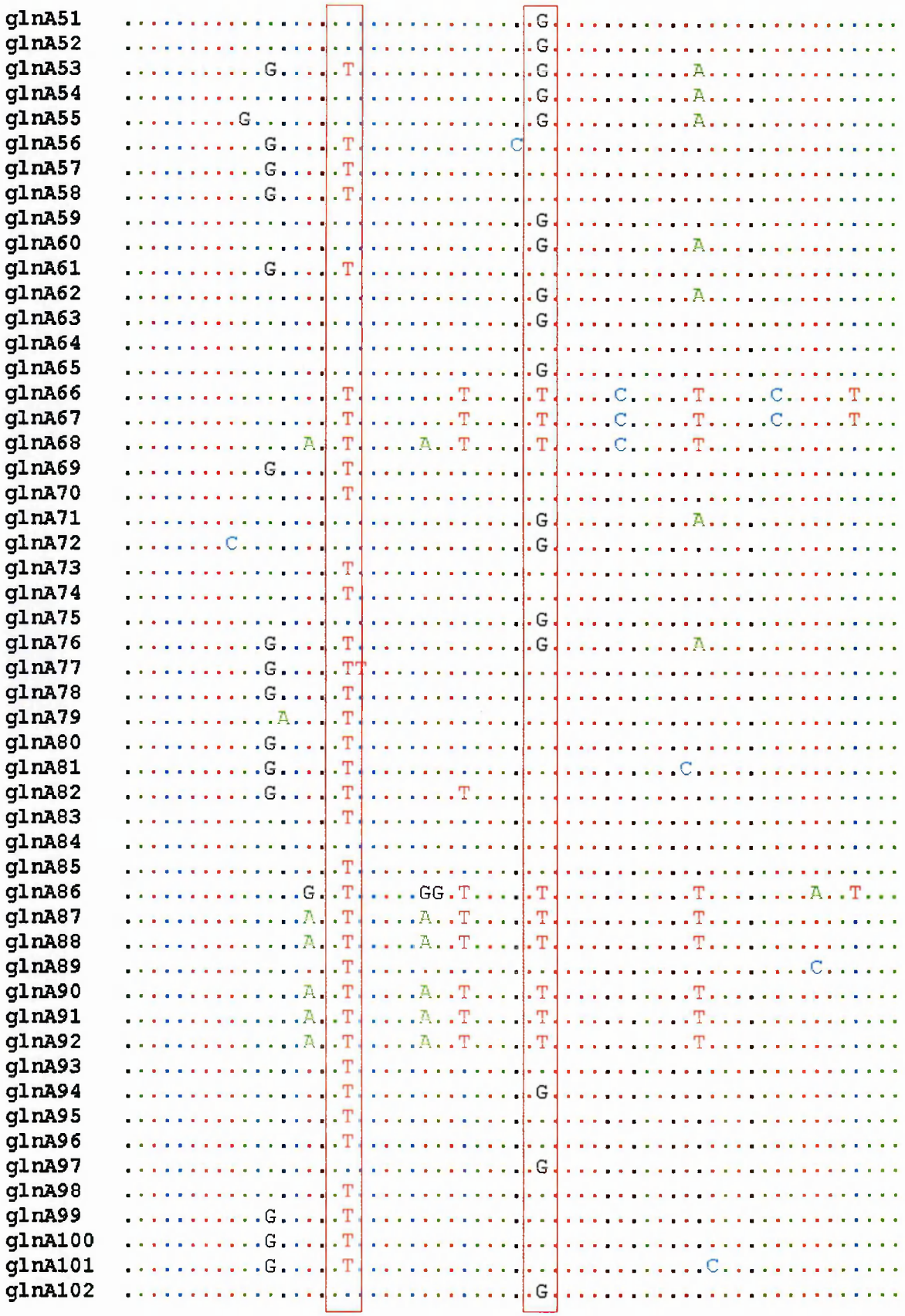
Appendix 2.9 *unc*A alignment showing SNPs at 333-339bp on *unc*A17




unc51 .....T.....T.....A..T..A.....
unc52 .....T.....T.....A..T..A.....
unc53 .....T.....T.....A..T..A.....
unc54 .....T.....T.....A..T..A.....
uncA55 .....T.....T.....A..T..A.....
uncA56 .....A.....T.....A..T.....T.....
uncA57 .....T.....T.....A..T..A.....
uncA58 .....T.....T.....A..T..A.....
uncA59 .....T.....T.....A..T..A.....
uncA60 .....T.....T.....A..T..A.....
uncA61 .....T.....T.....A..T..A.....
uncA62 .....T.....T.....A..T..A.....
uncA63 .....T.....T.....A..T..A.....

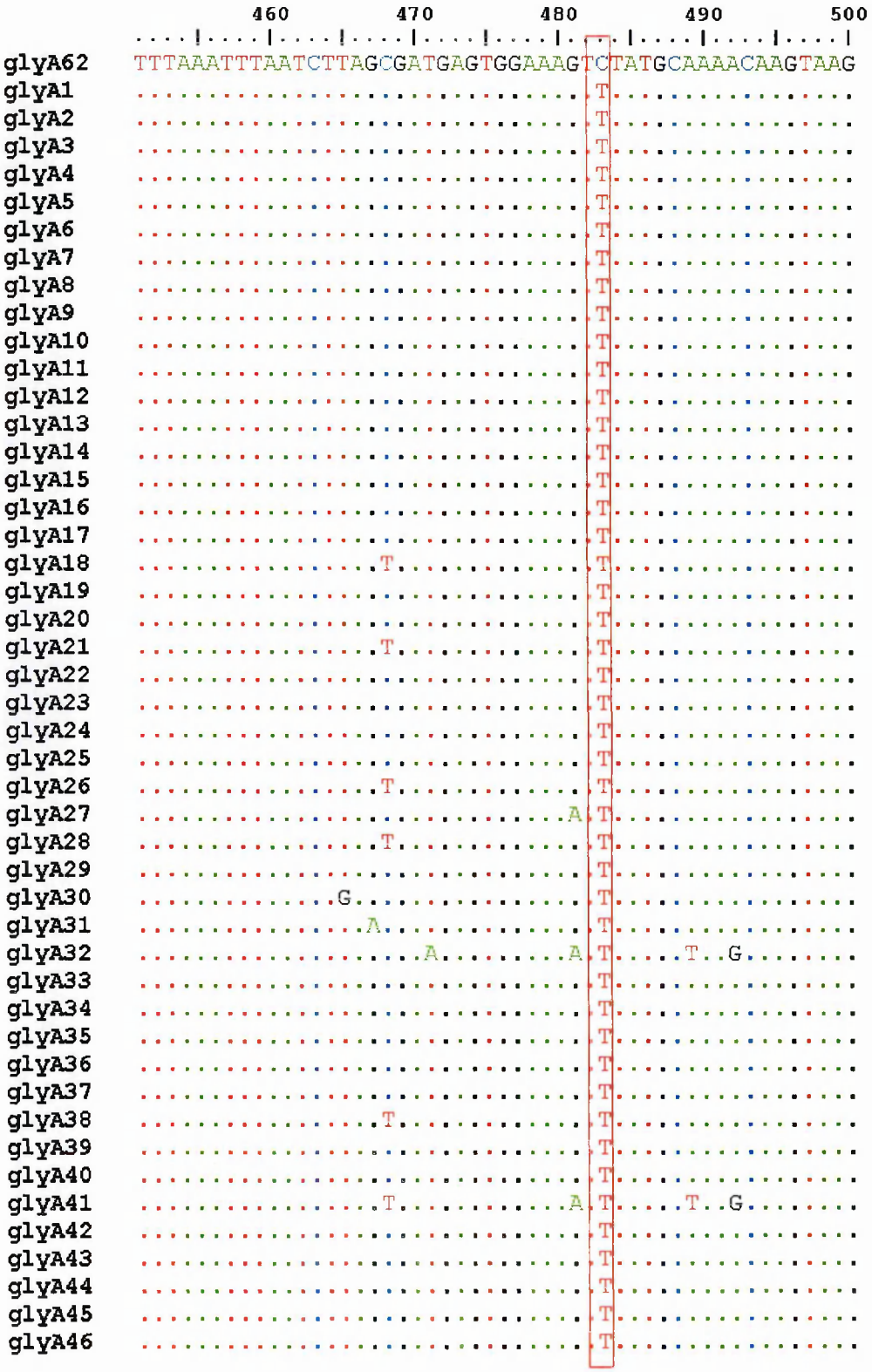
Appendix 2.10 *glnA* alignment showing SNPs at 18 and 33bp on *glnA21*

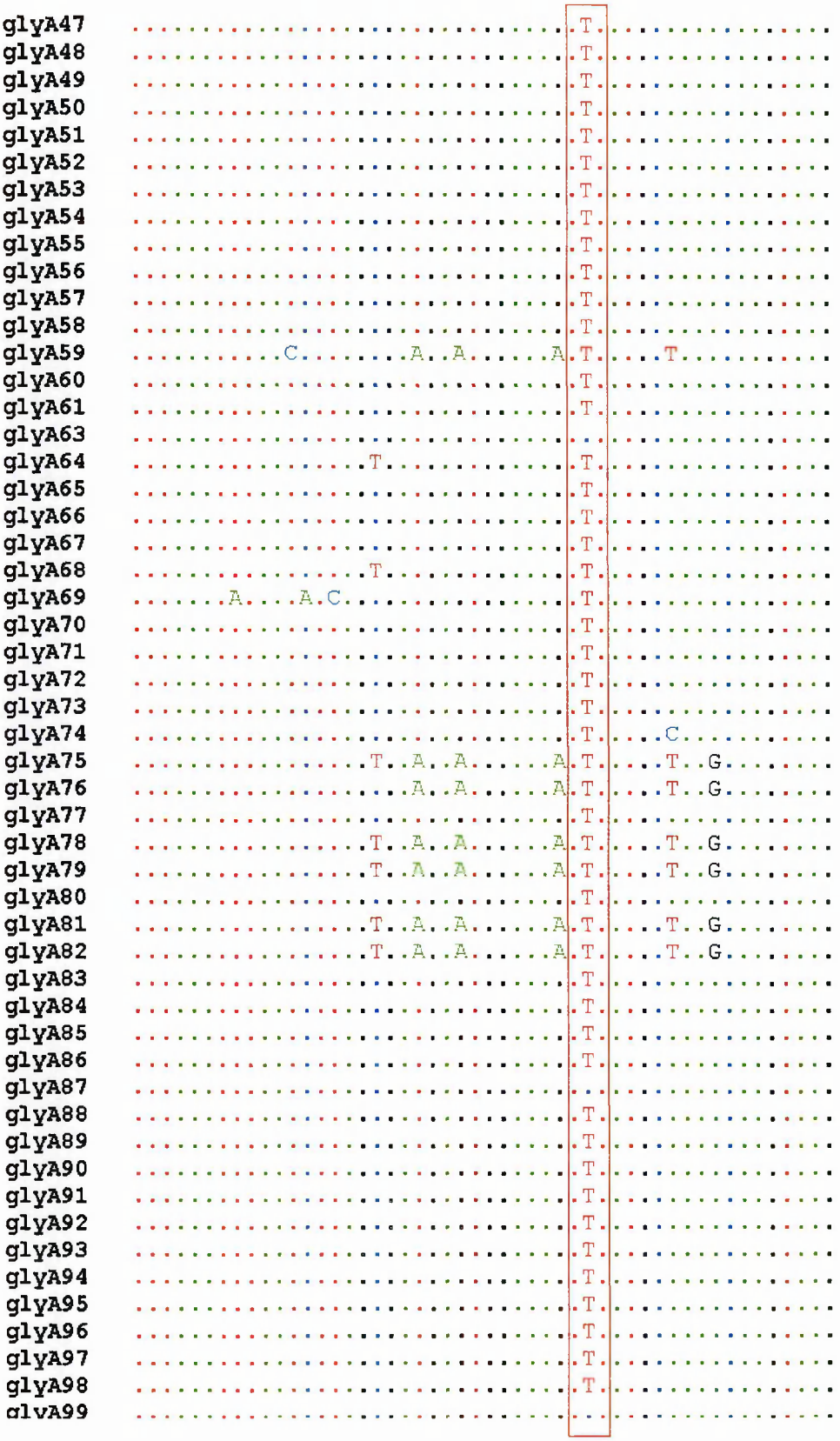


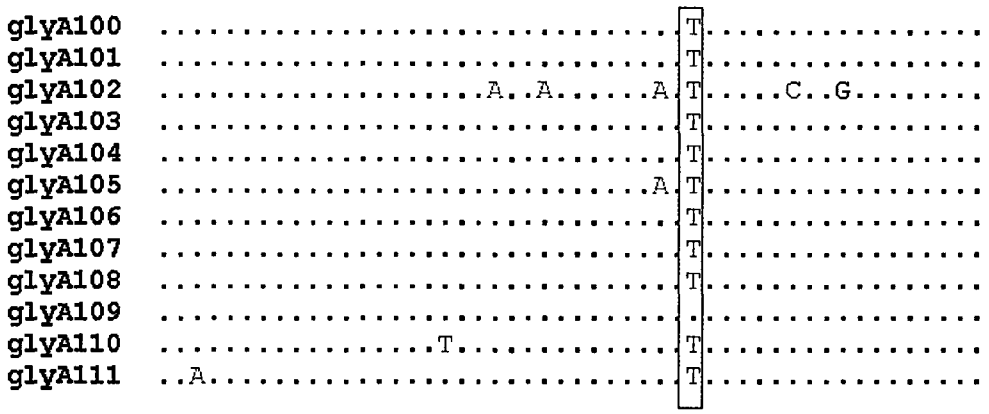

glnA51
glnA52
glnA53
glnA54
glnA55
glnA56
glnA57
glnA58
glnA59
glnA60
glnA61
glnA62
glnA63
glnA64
glnA65
glnA66
glnA67
glnA68
glnA69
glnA70
glnA71
glnA72
glnA73
glnA74
glnA75
glnA76
glnA77
glnA78
glnA79
glnA80
glnA81
glnA82
glnA83
glnA84
glnA85
glnA86
glnA87
glnA88
glnA89
glnA90
glnA91
glnA92
glnA93
glnA94
glnA95
glnA96
glnA97
glnA98
glnA99
glnA100
glnA101
glnA102


glnA103

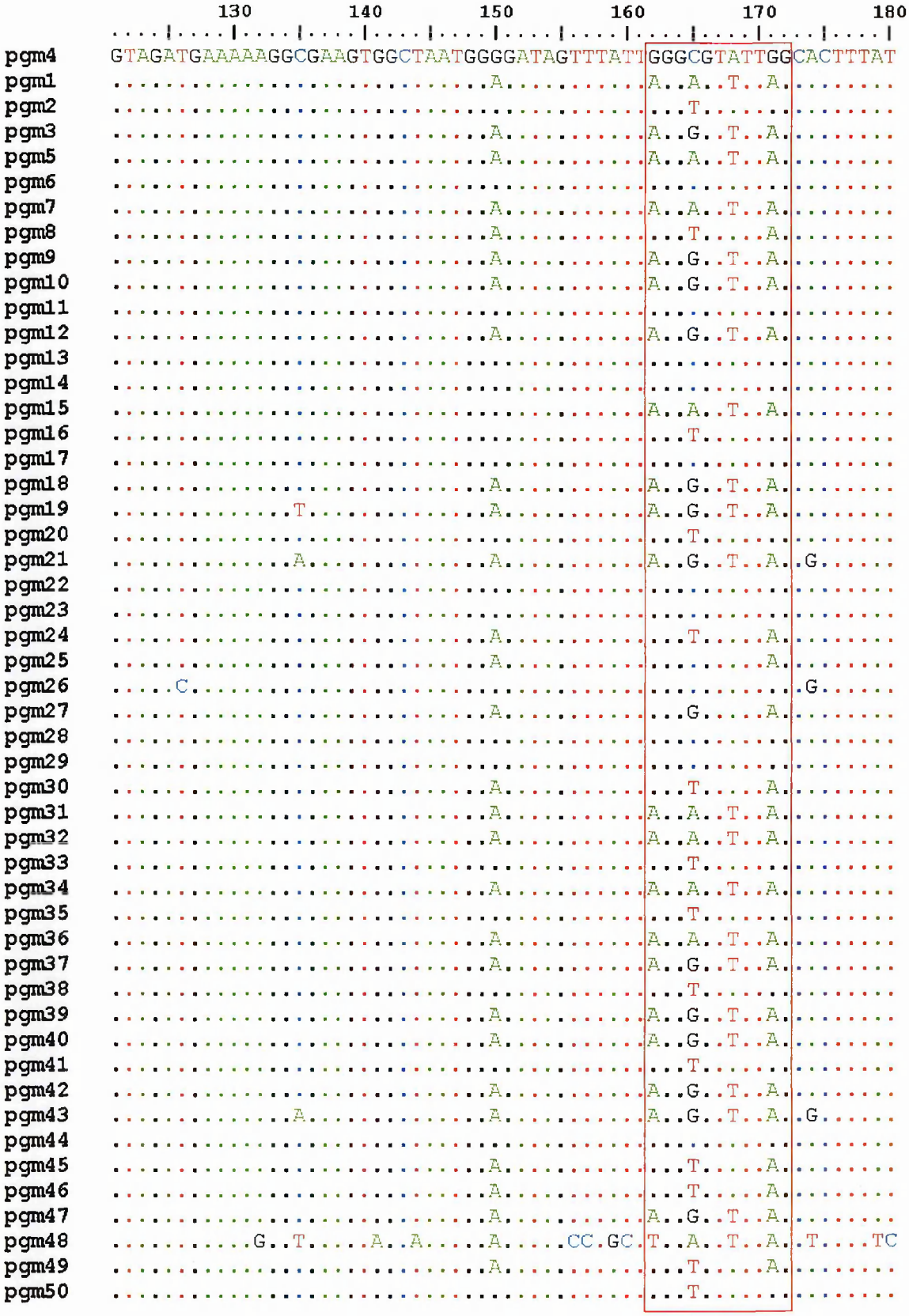
Appendix 2.11 *gly*A alignment showing SNP at 483bp on *gly*A62

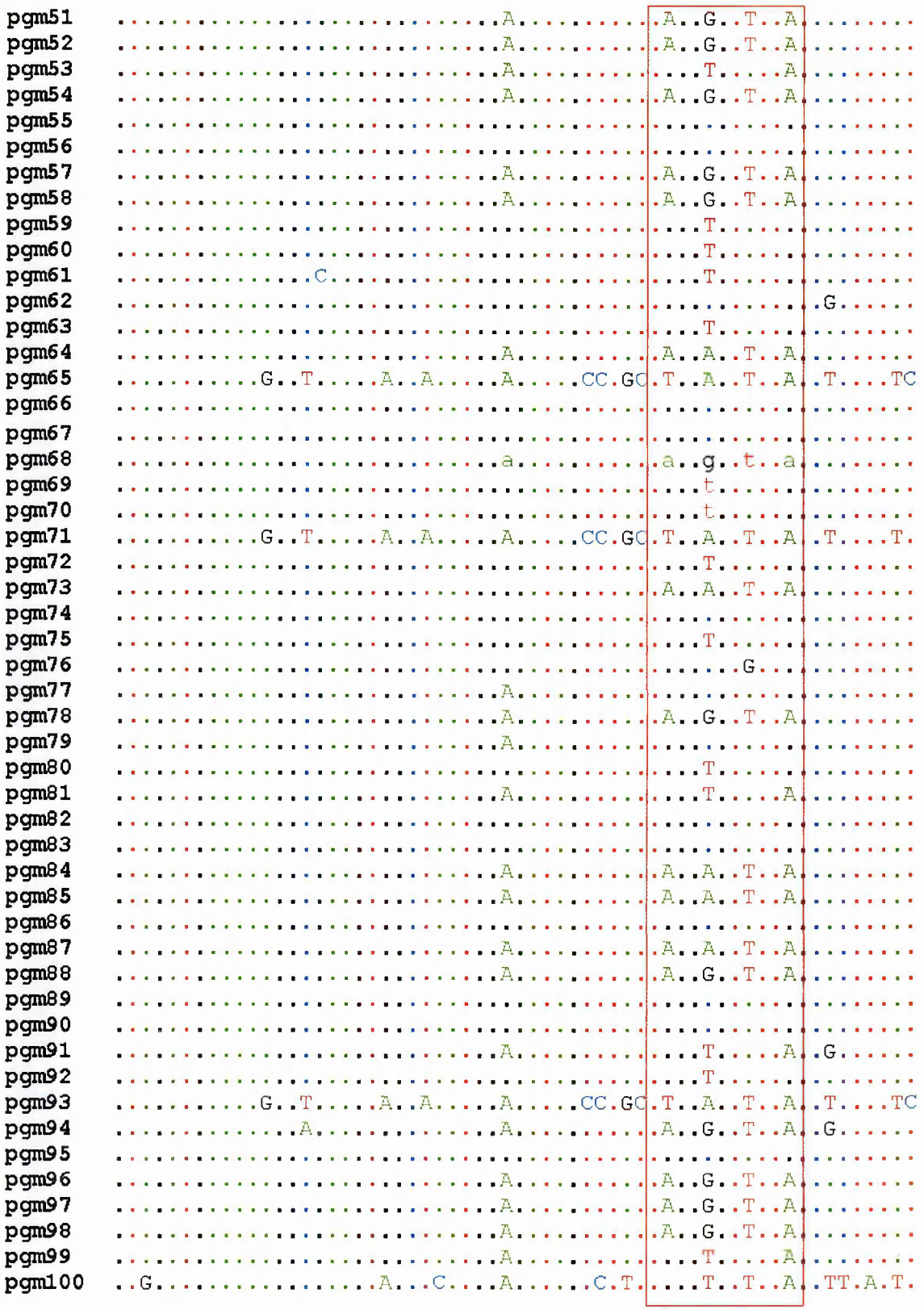



glyA47
glyA48
glyA49
glyA50
glyA51
glyA52
glyA53
glyA54
glyA55
glyA56
glyA57
glyA58
glyA59
glyA60
glyA61
glyA63
glyA64
glyA65
glyA66
glyA67
glyA68
glyA69
glyA70
glyA71
glyA72
glyA73
glyA74
glyA75
glyA76
glyA77
glyA78
glyA79
glyA80
glyA81
glyA82
glyA83
glyA84
glyA85
glyA86
glyA87
glyA88
glyA89
glyA90
glyA91
glyA92
glyA93
glyA94
glyA95
glyA96
glyA97
glyA98
glyA99


glyA100 ....T..............
glyA101 ....T..............
glyA102 ..................A..A......A.T.....C..G.....
glyA103 ....T..............
glyA104 ....T..............
glyA105 ............................A.T..............
glyA106 ....T..............
glyA107 ....T..............
glyA108 ....T..............
glyA109 ....T..............
glyA110 ................T.............T..............
glyA111 ..A...........................T..............

Appendix 2.12 *pgm_* alignment showing SNPs at 162-171bp on *pgm_4*




pgm51
pgm52
pgm53
pgm54
pgm55
pgm56
pgm57
pgm58
pgm59
pgm60
pgm61
pgm62
pgm63
pgm64
pgm65
pgm66
pgm67
pgm68
pgm69
pgm70
pgm71
pgm72
pgm73
pgm74
pgm75
pgm76
pgm77
pgm78
pgm79
pgm80
pgm81
pgm82
pgm83
pgm84
pgm85
pgm86
pgm87
pgm88
pgm89
pgm90
pgm91
pgm92
pgm93
pgm94
pgm95
pgm96
pgm97
pgm98
pgm99
pgm100